

West Scandinavian Vowel Systems  
and the Ordering of Phonological Rules

by

Stephen R. Anderson

B.S., Illinois Institute of Technology (1966)

Submitted in partial fulfillment  
of the requirements for the  
degree of Doctor of Philosophy

at the

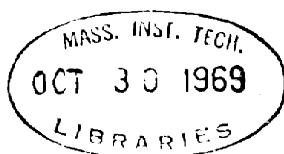
Massachusetts Institute of Technology

September, 1969

Signature of Author .....  
*U.S. Foreign LITERATURES & LINGUISTICS* Department of Modern Languages,  
18 Aug., 1969

Certified by .....  
Thesis Supervisor

Accepted by .....  
Archives Chairman, Departmental Committee  
on Graduate Students



# West Scandinavian Vowel Systems and the Ordering of Phonological Rules

Stephen Robert Anderson

Submitted to the Department of Foreign Literatures and Linguistics  
on 18 August, 1969, in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy.

The aim of this work is two-fold: first, to present an analysis of the more important aspects of the vowel phonology of Icelandic and Faroese; and secondly, to make use of these and other facts to present a theory of the organization of the phonological component of a generative grammar. This theory is based in large part on that presented by Chomsky and Halle in the Sound Pattern of English, differing from that work most importantly in allowing the total set of ordering requirements in the grammar to be inconsistent with any possible single linear arrangement.

Chapter I presents the basic notion of phonological rules and their interaction. The formal characteristics of a linear ordering are presented, and the alternative theory of local ordering to be defended in this work described.

Chapter II is an analysis of some of the more important problems of Icelandic phonology. The u-umlaut rule is considered in some detail, and it is concluded that its modern form is essentially the same as its form in Old Norse. The rules of vowel shift and stress shift in e-diphthongs are among the more important rules presented.

Chapter III is an analysis of the structure of the vowel system of modern Faroese. The large variety of diphthongs characteristic of the language are shown to be derivable from a simple system of underlying long vowels. The rule which inserts glides to break up vowel sequences is considered. The alternation between diphthongs and vowels plus palatal-velar consonants (the verschärfung) is shown to follow from otherwise-needed rules of the language.

Chapter IV discusses the role of the various notational conventions in the theory of phonology. The variables, braces, and parentheses notations are discussed in terms of their role as evaluatory functions and in terms of the distinction between conjunctive and disjunctive application. The neighborhood notation is discussed in some detail, and it is shown that this notation does not determine conjunctive or disjunctive ordering uniquely. An example is discussed which indicates the need for an extension of the principle of disjunctive notation beyond the fact of any particular formal similarity of rules.

Chapter V considers several examples that support the theory of local ordering as opposed to the theory of linear order. In the course of this discussion, several indeterminacies in the theory are uncovered, and partially resolved. The examples discussed are from Menomini, Kasem, Sanskrit, Turkish, Caddo and French. In connection with these last two,

Chafe's theory of 'persistent rules' is considered. Throughout the chapter, Kiparsky's notion of marked and unmarked order plays a significant role.

Chapter VI is a largely inconclusive and incomplete discussion of some of the implications of the theory of local ordering for other parts of linguistic theory. The consequences for Chomsky and Halle's theory of markedness are considered.

Thesis Supervisor: Morris Halle

Title: Professor of Linguistics

For Kitten, without whom it would never have been possible to complete this; and Thor, on account of whom it nearly wasn't.



## Table of Contents

	page
Abstract	ii
Dedication	iv
Table of Contents	v
Introduction	viii
 Chapter I - Basic Assumptions	
I.1 Phonological rules	1
I.2 The interaction of rules	4
I.3 Types of ordering relation	8
I.4 The organization of systems of rules	11
Footnotes to Chapter I	20
 Chapter II - The Icelandic Vowel System	
II.1 Inventory of vowels	24
II.2 The u-umlaut rule	29
II.3 The order of u-umlaut, and Old Norse skaldic verse	38
II.4 Sketch of the ablaut system	54
II.5 Breaking	59
II.6 The i-umlaut rule	64
II.7 The resulting vowel system	67
Footnotes to Chapter II	72
 Chapter III - The Vowel System of Faroese	
III.1 Introduction	76
III.2 Inventory of vowels	77
III.3 The glide-epenthesis rule	82
III.4 The <u>Verschärfung</u>	84
III.5 Some apparent exceptions	90
Footnotes to Chapter III	98
 Chapter IV - Notational Conventions	
IV.1 Evaluation procedures and notations	100
IV.2 Some exceptional types of ordering	109
IV.3 Iterated application and infinite schemata	116
IV.4 Neighborhood rules	122
IV.5 Disjunctive order and the status of notations	137
Footnotes to Chapter IV	148

Chapter V - The Principles of Rule Ordering	
V.1 Introduction	152
V.2 Menomini	153
V.3 Kasem	158
V.4 Grassmann's law in Sanskrit	164
V.5 Turkish harmony	178
V.6 Chafe's theory of 'persistent rules'	190
Footnotes to Chapter V	196
 Chapter VI - Local Ordering and Phonological Theory	
VI.1 Some implications	199
VI.2 The status of 'marking conventions'	201
Footnotes to Chapter VI	210
 Bibliography	211
 Biographical Note	218

My own subject, Phonetics, is one which is  
useless by itself, while at the same time  
it is the foundation of all study of language,  
whether theoretical or practical.

Henry Sweet

## Introduction

The title of this work, like most titles, is deficient in that it suggests both too much and too little. The use of the term 'West Scandinavian', in particular, is bad -- properly speaking, the West Scandinavian languages are all of those more or less closely derived from Old Norwegian, including Icelandic, Faroese, the dialects of Western Norway that form the basis of modern Landsmål Norwegian, and the dialects of Old Norse that were spoken in Shetland and in the Orkney Islands up until about the 18th century. These latter are virtually inaccessible to investigation at this point; there are really no clear written records of the language, and the dialects investigated by Jacob Jacobsen and others had lost a great many features through contact with Celtic and English. About the only materials available are contained in Jacobsen's dictionary, which is a compilation of the Norse-based words in the dialect of the 19th century inhabitants; these are very difficult of interpretation, and I have simply ignored Orkney and Shetland Norn in this work.

The exclusion of Norwegian is vastly less defensible, and I can only claim that to have included it would have involved vastly greater amounts of work without a great deal of light being shed on the problems of Icelandic and Faroese phonology which form the main part of the work. These two languages display many common features, but are sufficiently different to be treated quite independently. One might refer to them (as opposed to Norwegian) as 'insular West Scandinavian', but this would seem to do an injustice to the people of Shetland and Orkney. I must therefore simply state that my treatment of 'West Scandinavian' is restricted to Icelandic and Faroese, with a few glimpses of the Old Norse that both are related to.

In any case, my aim in presenting these discussions of West Scandinavian has been to provide a basis for the discussion and evaluation of a particular approach to the organization of a phonological description, in particular, of a theory of rule ordering. By discussing the rules whose application supports the ordering assumptions made here within the framework of the more or less complete description of a large portion of the phonologies of these two languages, I hope to have made the whole description (including the ordering arguments) more plausible. In the later chapters, in which the phonological theory is subjected to a more detailed examination, I have fallen back on the old technique of discussing a few rules in (artificial) isolation from the rest of the language; I hope that the fact that most of the examples chosen are virtual clichés of phonological description will make up for their disconnected appearance.

The structure of the work is rather similar to that of Chomsky and Halle's Sound Pattern of English, and the questions to which it is addressed are much the same as those that occupied that work. If much that is here seems to be nothing but an attack on claims made by Chomsky and Halle, this is perhaps inevitable, and I hope it will not be interpreted simply as iconoclastic. As virtually the first linguists in history to concern themselves with formulating their claims about the structure of language in such a way as to be empirically confirmable or disconfirmable, it is natural that much of what goes on in the field today is concerned with the evaluation of Chomsky and Halle's work. The points raised so summarily in chapter VI below perhaps indicate a new direction being followed in phonology today, one on which Chomsky and Halle have typically gotten the first start, but one in which other investigators have shown considerable

independence. Much of what appears in this book is devoted to what has gone before, but it all seems to be points that should be gotten clear before going on.

Much of the phonological theory of SPE is assumed here. Though some lip service is paid to the idea of motivating phonology from scratch, this is more an expository device than a serious attempt, and the reader completely unfamiliar with generative phonology will probably not fare too well. It should also be pointed out that the transcriptions used in discussing the various examples are not all consistent with one another; I have used, in general, the most common transcription for each language. The use of // brackets indicates either underlying forms or forms of intermediate status; that is, any form to which there are still rules that may apply. The [] brackets indicate a (more or less) surface phonetic form; one to which no more rules of any interest to the discussion apply. Forms not in any kind of brackets, but simply underlined, are generally orthographic.

By this point in a graduate career, one has invariably contracted large numbers of debts, financial and intellectual, which are supposed to be satisfied by the thesis. On the first score, I would like to express my gratitude to the staff of the Research Laboratory of Electronics, MIT, for the manifold little things that make an education possible. I would also like to thank someone for the fellowship aid I received from the National Institutes of Health and the National Defense Education Act, Title IV, but I am not sure whom. I can thank Susumu Kuno of the Harvard University Computation lab, Bruce Fraser of the Language Research Foundation, and S. Jay Keyser of Brandeis University for the help they have extended to me at

various points in my graduate career by research support carried out under contracts to Harvard University by the National Science Foundation and to Brandeis University by the United States Air Force, among others. I am especially grateful to the American Scandinavian Foundation for awarding me a Thor Thors fund grant which enabled me to spend much of the summer of 1968 in Iceland, a trip from which this thesis derived much. While in Iceland, the staffs of Háskólabókasafnið and Landsbókasafnið were extremely helpful.

One of the most exciting things about an MIT education is the opportunity for exchange of ideas with teachers and colleagues. Among these, I would like to mention Michael Brame, Wayles Browne, S. Jay Keyser, Einar Haugen, Ken Hale, Wayne O'Neil, David Perlmutter, H.A. Roe, Arthur Schwartz, J.F. Staal, and Nancy Woo. Roy Wright and his wife Stebba were extremely helpful with the Icelandic material. Though I have never met Hreinn Benediktsson, his excellent and comprehensive articles on various aspects of Icelandic phonology have provided a very important stimulus to much of my work. Paul Kiparsky was a primary source of ideas during the initial stages of work; had I been able to work with him during the last year, I am sure the thesis would have benefited greatly.

It is to Noam Chomsky, and especially, Morris Halle that this work (as well as its author) is most indebted, however. Both have given extremely generously of their time and their ideas, and by virtually creating the field, have made it possible for others to benefit. To Professor Halle, to whom I owe the most important part of my graduate education, the distinction between solutions that work and answers, this thesis owes most of its present form. The usual disclaimer at this point is at least as apposite

to this thesis as to any other: whatever merit there is in it is due largely to those named above, while its defects are squarely with its author.

A preliminary version of some of the material in chapter II appeared previously as Anderson, 1969; much of chapter III appeared as Anderson, 1968. The publication of section V.4 is planned for the near future.



## Chapter I - Basic Assumptions

I.1 Phonological rules. In setting out to characterize a speaker's knowledge of the sound system of his language, we are led to consider each utterance as made up of a sequence of discrete units.<sup>1</sup> As has often been noted, it is not in general possible to identify these units directly and uniquely in the physical description of the speech signal, but the positing of such theoretical entities has been the basis of nearly all of the productive results of modern linguistics in the field of sound structure, and their use is thereby justified to the extent the conclusions based on it are accepted as valid insights into the structure of language.

It would be possible, in principle, for a speaker's knowledge of the sound structure of the elements of his language to be characterized simply by a gargantuan list of all such elements that could possibly occur in utterances in the language. While it would be impossible so to characterize syntactic structure in light of the (recursive and therefore) unlimited nature of such knowledge, a language contains a finite number of distinct lexical items, and these could possibly be memorized as a list. Were this a true description, phonologists could simply give up and go to the beach, but in fact speakers appear to reduce the burden of memory by memorizing only the properties of an element of the list which are necessary to distinguish it from other elements, and generalizing properties of the entire list. These general properties are then assumed to be part of each member of the list, unless one of the distinguishing properties of an element is precisely the fact that some otherwise valid generalization does not extend to this item.

Such generalizations about the elements of the language are fundamentally of the form:

- 1.1 If a linguistic form has properties  $x_1, x_2, \dots, x_m$ ,  
then it also has the properties  $y_1, y_2, \dots, y_n$ .

The psychological and behavioral evidence for such generalization, while somewhat nebulous, is quite convincing. Thus speaker behavior does not in general have the random and unstructured character of lists. When speakers learn new forms, these are not unrelated to the ones they already know, but rather have properties of the original list generalized to them. When they make mistakes, these do not consist in the substitution of completely unrelated forms, but rather in the substitution of forms with very close connections in terms of the systematic properties of the set of forms of the language. When languages change, a very frequent form of such change is the extension of already existing generalities to include new sets of forms not hitherto subject to them. Such kinds of 'evidence' could be prolonged indefinitely, and all such behavior is inconsistent with a denial of reality to rule-like statements about regularities of language.

When we further investigate the nature of such regularities, we see that they are in general statable not in terms of segments directly, but rather in terms of classes of segments, such as stops, nasals, vowels, etc. We are thus led to represent the segments themselves not as unanalyzed units, but rather as collections of properties, each such property representing the status of the segment with regard to membership in one of the linguistically relevant classes. Investigation of these features leads to the further conclusion that they are in large part, though not entirely, basable on articulatory considerations; the precise basis of each feature, and the

specification of a set of such features which is adequate for the description of all natural languages is a task for a universal phonetic theory. While some consideration will be given in chapter VI below to the implications of a particular choice of a set of features (a matter of fundamental importance), it will be assumed here that such a set has already been chosen. This set will be largely identical with the features described and used in Chomsky and Halle 1968<sup>2</sup>, though with certain modifications that have been suggested since then.<sup>3</sup>

Returning to the nature of the rules that express the general properties of a language, we see that there is no reason, in principle, why the set of features whose values are specified by the properties  $y_1, y_2, \dots, y_n$  in a statement like 1.1 above should be disjoint from those whose values are specified by  $x_1, x_2, \dots, x_m$ . Thus, the rules of a language may not only fill in unspecified properties of a segment (as is basically assumed in traditional taxonomic phonemic theory<sup>4</sup>), but may also change some already specified properties. We are thus led to assume that the level of representation which is directly specified in the speaker's 'list' is a list of the properties which the individual unit would have, among others, if no other fact of the language intervened. Thus, independently necessary properties of the set of phonetic forms of the language (and only such properties) can lead to an item's representation by values which it does not in fact manifest phonetically. Extensive discussion of the relation between the properties of a phonetic form and those of its ideal representation will be found in work by Postal (1968) and Kiparsky (1968a). It should, however, be pointed out here that an important claim of phonological theory is that the sets of features needed for the two representations are the same.<sup>5</sup> We will refer

to the former of the two as the (systematic) phonetic representation of the form, and to the latter as its (systematic) phonemic representation. It is assumed that no intervening level has any privileged status.<sup>6</sup>

I.2 The interaction of rules. Let us now consider some examples of phonological rules in natural languages. In the Ural-Altaic languages, vowels are usually divisible into classes, such that the vowels of any given word are all members of the same class. The vowels of affixes usually differ depending on the class from which the vowels of the base are drawn. In the Turkic languages, words with back vowels are followed by forms of suffixes containing back vowels, and words with front vowels by forms of the same suffixes containing the corresponding front vowels. We could say then that such languages contain a rule of vowel harmony:

- 2.1 If the vowels of a base are front (resp:back), then the vowels of suffixes attached to that base are also front (resp:back).

In East Turki, however, there is a further wrinkle (cf. Poppe, 1965): there is no back vowel corresponding to i, and furthermore, the vowel a is replaced by e if the next vowel in the word (which may be either in the base or in a suffix) is i. This language, then, contains a different rule of vowel harmony:

- 2.2 If the first vowel of a word is front (resp:back), then following vowels which are not both high and non-round (i.e., not i) are front (resp: back).

and in addition, contains a rule of umlaut:

- 2.3 If a vowel is non-high and non-round (i.e., a or e), and the next vowel in the word is i, then it is front (i.e., e).

Now observe what happens when these generalizations are applied to the forms of the language. Consider the stem al- 'to take'. A verbal noun can be formed from this stem by the addition of the suffix -iś. When we do this, however, the umlaut rule applies (2.3), and we obtain eliś rather than \*aliś 'the taking'. By adding a passivizing suffix -in to the root al-, we obtain elin 'to be taken'. We can further add the nominalizing suffix -maq/-meq to this latter, and when we do, we obtain elinmaq 'to be taken (nominal)'. Note that in this form, the vowel of the suffix does not agree with the (phonetic) vowels of the base, as the vowel harmony rule (2.2) would seem to require, but rather with the ideal form the stem vowel would have had if the umlaut rule (2.3) had not affected the form. We say thus that while the base of elinmaq does not contain back vowels in its phonetic representation, its underlying or phonemic representation does contain back vowels, and it is these vowels that determine the form of the suffix, despite the fact that they are changed to front vowels by another rule of the language.

For another example, consider the vowel system of Telugu, as described in Bright (1966) and Kelley (1963). Telugu vowels are basically either 'tense' or 'lax', where 'lax' vowels are articulated at a lower tongue position than 'tense' vowels.<sup>7</sup> These vowels are distributed in the following way: if the final vowel of the stem is a (whose tense form is ə), all the other vowels are lax. Otherwise they are tense. Thus, we have the root [gu:ḍu] 'nest' with tense vowels, but [gU:ḍa] 'basket' with lax vowels. The addition of suffixes may cause the final vowel of the stem to disappear, but the vowel quality of the word does not change: e.g., the plural of [gu:ḍu] is [gu:ḷḷu] (from underlying /gu:ḍu+lu/), while that of [gU:ḍa] is [gU:ḷḷU] (from /gU:ḍa+lu/). Similarly, we find [kəmpu] 'a smell', and [kampa] 'a branch', but [kəmpa:] 'is it a smell?' and [kampa:] 'is it a branch?' (from

/kəmpu+a:/ and /kampa+a:/, respectively). In forms like these, it is apparently the vowel of the first syllable that distinguishes, say, [kəmpa:] from [kampa:], but when we consider the underlying representation of the forms, we see that this difference is a reflex of the difference in the final vowels of the stems, even though these are not directly manifested phonetically in the surface forms.

For a slightly different kind of example, consider the process of palatalization in Pame, an Otomean language of Mexico (for details, cf. Gibson 1956). In this language, when a morpheme whose underlying form begins with a cluster containing dental consonants (i.e., t, d, c=t̪s, s, n, or l) is placed after a morpheme ending in i, the dental consonants are changed to the corresponding members of the 'velar-palatal' series (i.e., k<sup>y</sup>, g<sup>y</sup>, č, š, ŋ<sup>y</sup>, and l<sup>y</sup>).<sup>8</sup> Some examples of this change are: /ski+t̪əhan?/ becomes [skik<sup>y</sup>əhan?] 'your soap'; /ni+sās/ becomes [nišās] 'you played (music)'; /ki+n̪ə/ becomes [kiŋ<sup>y</sup>ə] 'your tongue', etc. Consider some forms which contain more complex clusters: /ri+n̪thôit̪/ becomes [riŋ<sup>y</sup>k<sup>y</sup>hōik<sup>y</sup>] 'women'; /rómmadnt+i/ becomes /rómmaidnt/ by a metathesis which is not otherwise relevant here, and this in turn becomes [rómmaig<sup>y</sup>ŋ<sup>y</sup>k<sup>y</sup>] 'our jiggers (dual incl.)'. We see by these latter forms that all of the members of a cluster are affected by the palatalization process. But now consider /ki+n̪cáo/, which becomes [kinčáo] 'you burn yourself', and other forms containing clusters of n+š or č. Here the s or c becomes š or č normally, but the n is apparently not affected. We saw above in [rómmaig<sup>y</sup>ŋ<sup>y</sup>k<sup>y</sup>] that n is usually affected in clusters.

In Pame, as in most languages, nasal consonants are affected by a process of assimilation to neighboring obstruents. Thus, /n+phêit̪/ becomes [mphêik<sup>y</sup>] 'pigs', /n+tógn/ becomes [ndógn] 'flowers', /n+kwá?a/ becomes

[ŋgwá?a] 'your, his huge tamales (containing eggs or whole chicken)', etc. Here the initial nasals are assimilated to the point of articulation of the following stops. In /tiné?ep+t?n/, which becomes [tiŋ<sup>y</sup>é?ep?m] 'we appease him (du. pl. excl.)', and /kisá?p+kn/, which becomes /kisá?p+n/ by a cluster simplification rule which is not otherwise relevant here and then [kisá?pm] 'you(pl.) teach him', the nasal is assimilated to a preceding stop. Thus, Pame is seen to contain a rule whereby the point of articulation features of a nasal are specified as the same as those in a neighboring consonant on either side.

When we consider this process, together with the feature structure of c and s, we have an explanation for forms like [kinčáo] above, with apparently unpalatalized nasal. Unlike the other members of the 'velar-palatal' series, these two are articulated as palatal-alveolars (i.e., as [-ant, +cor, +high, -back]<sup>9</sup>) rather than as palatalized velars (i.e., as [-ant, -cor, +high, -back]). Thus, the nasal corresponding to c is n<sup>(y)</sup>, rather than ŋ<sup>y</sup>. The correct value of the nasal in [kinčáo] is thus specified by the nasal assimilation rule. But notice that if this is to be the case, the nasal assimilation rule must apply to the output of the palatalization rule. This is so because the palatalization rule applies to nasals (e.g., in [kiŋ<sup>y</sup>a], from /ki+na/). Thus, from underlying /ki+nčáo/ we will get /kiŋ<sup>y</sup>čáo/. It is only at this point that the nasal assimilation rule can apply to give the correct output [kinčáo]. Observe that it would not be sufficient to prevent the palatalization rule from applying to nasals if they appear in clusters; even if this step were taken, it would be necessary to apply the nasal assimilation rule to the output of the palatalization rule in order to get [rómmaig<sup>y</sup>ŋ<sup>y</sup>k<sup>y</sup>].<sup>10</sup>

In the examples from East Turki and Telugu, it would have been

possible to explain the results of the interaction of the pairs of rules involved by saying that in each case, both rules apply to the underlying form, by some simultaneous application algorithm. In the Pame case, however, we see that in some cases of rule interaction it is necessary to allow one or more of the rules to apply to representations that are not created until after other rules have applied. In light of this, we might propose that one starts with an underlying form, applies to it all rules that can be applied, and then applies all rules that can be applied to the output from this process, and so on until no more rules are applicable. We have already discussed cases, however, where this procedure would give the wrong results. On this assumption, for instance, a Telugu form like /kəmpu+a:/ would, after being affected by the external sandhi rules (giving /kəmpa:/) be subject again to the rule specifying the tenseness of vowels in terms of the final. This would convert /kəmpa:/ into the incorrect \*[kampa:], rather than the correct [kəmpa:]. Clearly, if rules are to apply to the output of other rules, they must not be allowed to do so indiscriminately. In this case, for instance, we must say that the grammar of Telugu is such that the laxing rule must apply before and not after the rule which syncopates the final vowel of the root before endings such as the plural, etc. We have thus established a basis for including ordering relationships among the rules of a grammar: when two rules apply to the same form, one or the other takes precedence. This notion is reflected formally in the theory of grammars by performing the operation corresponding to the one before that corresponding to the other, rather than vice versa.

I.3 Types of ordering relations. As we have seen above, the rules of the grammar must apply in a (partial) order. Rules tend to apply in as many



forms as possible; obviously, the place of a rule in the ordering relative to other rules can affect the number of chances it will get to apply. We can then say that some orderings are more natural than others, in the sense that they permit one of the rules to apply in more cases than would be true if they applied in another order. To capture this intuition within linguistic theory, Kiparsky (1968b) has suggested a distinction between marked and unmarked orderings of pairs of rules, depending on the generality of applicability they permit. This distinction, which is related to other issues of the complexity of grammars which will be discussed in chapter IV below, will now be briefly described.

Given two rules, it is of course possible that they apply to totally disjoint classes of forms, and hence fail to impinge on one another in any way. In such cases, it is not possible to describe an ordering relation between them as marked or unmarked, since neither affects (ex hypothesi) the extent of application of the other. If such is not the case, however, the following situation may obtain: one of the two rules  $R_1$  and  $R_2$  (say  $R_1$ ) has the effect of producing some instances of forms to which the other ( $R_2$ ) could apply. Thus, if the rules apply in the order  $R_1 - R_2$ , then  $R_2$  will be able to apply in more cases (namely, those produced by the application of  $R_1$ ) than if they applied in the order  $R_2 - R_1$ . If this is the case, we say that the order  $R_1 - R_2$  is a feeding order, and in conformance with the defining principle above, we say that if a feeding order is possible between  $R_1$  and  $R_2$ , it is the unmarked or more general condition for this order to obtain. Conversely, if the rules must apply in the other order, we say that this is the marked, or less general condition. Kiparsky has provided support for this notion by showing that in historical change it is fairly common for rules whose ordering at one point in time contradicts a feeding

order (i.e., stand in a marked order with respect to one another) to be reordered at a later point so as to establish such an unmarked order, but that the other type of conceivable change (i.e., from unmarked to marked order) is highly unlikely, if it is possible at all. For an example of rules which stand in a feeding order to one another, recall the rules of palatalization and nasal assimilation in Pame. Here the palatalization rule creates new instances of nasal plus non-homorganic stop for the assimilation rule to apply to; hence, the order palatalization - assimilation is a feeding order. In East Turki, however, the rules of vowel harmony and umlaut are ordered so as to violate a possible feeding relationship; after the umlaut rule applies, the vowel harmony rule could change the quality of the suffix vowels, but does not. The harmony rule thus applies before the umlaut rule, a situation which we would expect to be vulnerable to historical change because it involves a marked ordering.

Another possible relationship between  $R_1$  and  $R_2$  is the following: suppose that, instead of creating a class of new representations to which  $R_2$  could apply,  $R_1$  has the effect of altering some forms so that, while they were originally subject to  $R_2$ , they are no longer subject to it after the application of  $R_1$ . Such an ordering relationship we can call a bleeding order. It will readily be seen that the notion of markedness is applicable to such a case also: but in this case it is the marked condition for a bleeding order to obtain, while the unmarked case is for a bleeding order to be avoided. This conclusion is also supported by evidence from historical change adduced by Kiparsky. We will see below that the applicability of the notion of markedness to orderings is not restricted to cases describable in terms of fixed orders of pairs of rules, independently of other aspects of the grammar, but the notions of feeding and bleeding order provide a useful

first approximation to the notion.

The fact that rules conform to a principle of maximal generality of application, while an empirical result supported by empirical evidence, is not wholly unexpected. If, as is assumed here, the same set of features is used to characterize both phonemic and phonetic representations, it follows that the representations which result from the application of rules have the same sort of status as those that serve as input to the rules. There is thus no reason why these outputs should not in turn be subject to all of the other rules; if it is necessary to impose an ordering on two rules which prevents this, such as a bleeding order or a non-feeding order, it should count as an extra complexity in the grammar. Notice that the crucial assumption here is that of homogeneity of representation; if this were not the case, and the output of a rule had a different sort of status from its input, it would be unexpected and unusual if the output of any rule were in turn subject to other rules, and such an ordering would be expected to be costly. This would be the case in a theory in which phonemic and phonetic representations were completely distinct, as in Lamb's phonemic theory among others (cf. Lamb, 1966). The fact that natural languages appear to involve numerous orderings, and that these ordering relations tend to conform to a principle of maximal applicability (as suggested by Kiparsky) is, indirectly, evidence for the correctness of a homogeneous system of representation for both the phonemic and the phonetic levels.

I.4 The organization of systems of rules. Given an ordering relation between rules, it is natural to enquire further into its properties; i.e., into the organization of such pairwise ordering relationships into a principle of application of the full set of rules which express the complete

range of generalizations about the phonological structure of the language. This question will be considered in detail in chapter V below. In order to establish a framework for the discussion of West Scandinavian phonologies in chapters II and III below, however, it is necessary to say a bit about these issues here.

It has already been observed that the full set of rules must apply sequentially (at least in part); that is, that it would not be possible for all of the rules to apply simultaneously to a single level of representation, producing their respective effects with no reference to one another. One might then suggest that the rules apply in 'random sequential' fashion; that is, that each rule applies at any point where the conditions for its operation are met. Examples already discussed showed, however, that in some cases rules can apply only in certain positions with respect to other rules in the total sequence of applications; thus, it is necessary to include in the grammar of a language at least some restrictions on the order of application of the rules.

One might then plausibly suggest that the set of rules be arranged in a single sequence, and then applied one at a time starting at the beginning of the list and proceeding toward the end, with each rule applying to the results of the application of the preceding rule, and continuing to the end of the list, without re-applying any rule which has already been passed over in the list. Such a principle of linear ordering has been the basis of much of the work in generative phonology in the last ten years or so, as well as other traditional work such as Bloomfield (1939). Examples discussed by Chomsky (1967) and Chomsky and Halle (1968) as well as numerous other papers lead to a modification of the linear ordering principle along the following lines: the rules are applied in linear sequence, except that if

two adjacent rules are related in certain ways, the applicability of the first precludes the application of the second, and this latter rule is thus skipped in the sequence when the former applies. It will be seen in chapter IV below that this principle of disjunctive ordering is invoked if the rules are formally related by the variable or parentheses notations,<sup>11</sup> and in at least some cases of rules which are related by the principle of "neighborhood rules".<sup>12</sup> It will be noted that while these formal devices are not entirely adequate to specify the class of disjunctively ordered rules, they constitute a very close approximation.

Disjunctively ordered rules, then, constitute a class of counter-examples to the principle of linear ordering, but do they in fact constitute the only such class? It has generally been assumed that the relation of order among phonological rules has a number of properties in common with other, more familiar ordering relations, such as the relations 'greater than' and 'less than' holding between natural numbers. Important properties of such relations include those listed in 4.1 below:

- 4.1 Given a relation R holding between the members  
a, b, c, ... of a set S:
- a. R is assymetric: that is, if (aRb), then it is not also the case that (bRa).
  - b. R is irreflexive: that is, there is no element a such that (aRa).
  - c. R is transitive: that is, if (aRb) and (bRc), then (aRc).

In terms of the rules of a grammar and the relation 'precedes', these properties can be translated as:

- 4.2 a. Assymmetric: No rule both precedes and follows<sup>13</sup>  
the same rule.
- b. Irreflexive: No rule precedes itself.
- c. Transitive: Given the three rules  $R_1$ ,  $R_2$ , and  $R_3$ ,  
then if  $R_1$  precedes  $R_2$ , and  $R_2$  precedes  $R_3$ , then  $R_1$   
must also precede  $R_3$ .

If the relation of order among phonological rules meets these conditions, which will be referred to below as the requirement of linearizability, then it can easily be shown that the set of (pairwise) ordering relations which must be defined on a set of rules will always be consistent with at least one linear ordering of the rules. Exceptions to the principle of linear ordering will always, in that case, be of the sort captured by the principle of disjunctive ordering: that is, exceptional facts about the applicability of the rules, which do not, however, affect their order of application.

Now while the properties 4.2 seem to be a natural enough set of assumptions about the ordering relation in phonology, it must be emphasized that they are assumptions, and as such subject to empirical confirmation or disconfirmation. While the solutions that have been provided for particular problems within the framework of generative phonology have in general been consistent with such a set of assumptions, it is still possible that further evidence would lead us to reject one or all of them. In chapter IV below, for example, it will be suggested that at least some of the examples dealt with in terms of an infinite schema in the framework of SPE should, in fact, be treated as examples of rules which apply iteratively to their own output. Such a principle is in direct contradiction to 4.2b, of course, and if supported would lead us to abandon it. The damage done to the spirit

of the linearizability requirement by the overthrow of 4.2b alone is minimal; it simply allows a rule to appear more than once on the list, but still requires all such appearances of the rule to constitute a sequence and to be ordered in the same way with respect to the other rules of the grammar. Obviously, 4.2a and 4.2c are the heart of the principle.

Apparent counterexamples to 4.2a or 4.2c ("ordering paradoxes") have arisen in work on a variety of languages. Many of these have been dealt with by reformulations of the rules, redefinitions of features, etc.: devices which permit the requirement of linearizability to be maintained, though at the cost of counter-intuitive claims about the structure of the languages in question. Some such cases will be reanalyzed in chapter V below, where it will be shown that the correct solutions to these problems do in fact constitute counterexamples to 4.2. The existence of such cases, then, will constitute positive counter-evidence to the principle of linear ordering, and as such will require us to conclude that the relation of ordering which is relevant for phonology is not, in fact, analogous to the ordering relations holding between natural numbers, etc. As such, it will be necessary to introduce other principles of application of rules to deal with the so-called paradoxes.

Attempts have already been made, of course, within the theory of generative phonology, to introduce such other principles. The most extensively investigated proposal of the sort is the principle of the transformational cycle, whereby the rules are organized linearly, and apply in a linear order first to the innermost constituents of the syntactic bracketing,<sup>14</sup> then reapplying the entire list of rules to the next highest level, etc., until the highest order of brackets is reached. Such a principle has found its most extensive support in the analysis of English

stress begun by Chomsky, Halle and Lukoff (1956), and further reported in SPE and in Halle and Keyser (forthcoming). While numerous efforts have been made to apply this principle to segmental phonology, the analyses predicated on this assumption have generally been shown to be incorrect in one way or another. Since stress assignment rules, which are in general more intimately connected with syntactic structure than most segmental rules, will not be considered in detail in this study, it is safe to assert that the principle of the transformational cycle has, at this point, no substantive support in the domain under investigation here. Recent work by M. Brame, which has not yet been published, suggests a similar principle may be necessary to explain the facts of Arabic phonology, but until these results are more conclusive, it is impossible to consider their implications.

It is necessary here to say something about the principles which will be assumed to underlie the application of a set of rules for the purposes of chapters II and III. These assumptions will be examined in more detail in chapter V below. We assume that ordering relations, just like other properties of rules, are stated in the grammars of particular languages in precisely those cases where they are not predictable by general principles. Thus, the grammar of a particular language will contain a statement about the order of application of two rules only in case that order is both necessary and unpredictable. The rules may well be such that it is unnecessary to establish any order of application for certain pairs, and if so there will be no ordering defined between these rules. This is in general the case if the applications of the two rules do not impinge on one another. Such a situation is not, of course, possible in a linearly ordered grammar, where each rule's position on the list defines its position relative to all of the other rules on the list, but it seems intuitively correct.



It has been mentioned that some orderings are not to be stated, but rather predicted. It is desirable to utilize the insights into the differing naturalness of different possible orders of pairs of rules in defining the general principles of rule orderings. It will therefore be assumed that the sequence of application of the rules is defined, where possible, by the unmarked order of the rules in question, and that rules that apply in this order will contain no ordering specification in the grammar. Rules which must apply in a marked order, on the other hand, as well as rules whose order is important, but for which no unmarked order exists, must be related by an explicit statement in the grammar to this effect. Consistent with Kiparsky's observations about the direction of historical change, we can say that a possible form of simplification of grammars is the loss of an ordering restriction, a change which becomes quite parallel to the loss of a feature or features from a rule, which is a well established form of generalization over time. Both types of change result in more general, less marked grammars.

It would seem that this proposal is notationally equivalent to a suggestion that all of the relevant orderings are defined, and a principle of evaluation then applies to the grammar which considers some of these (the 'unmarked' ones) less expensive than others. It will be seen below, however, that the notion of 'unmarked order' as formulated so far contains certain fundamental indeterminacies whose resolution will lead to empirical differences between the proposals.

The following, then, is approximately the principle of rule application which is assumed here: given an underlying form in the language, which contains the rules  $R_1, R_2, \dots, R_n$  together with a collection of ordering relations of the form " $R_i$  must precede  $R_j$ " (or, equivalently, " $R_i$  must not

follow  $R_j$ "), consider all possible sequences of the rules. Discard any such possible sequence which contains two rules whose order contradicts one of the restrictions. Of the remaining possible orderings, discard any which contains two rules in a marked order unless this order is explicitly provided for in one of the restrictions. The result of this procedure may not, in general, be a single ordering, but it is necessary to impose, as a general global condition of adequacy for grammars, the requirement that all such possible orders remaining after the eliminations described above yield the same result. Such a condition is clearly satisfiable for any grammar; in the limiting case it may be necessary to impose an order between every pair of rules, in which case the eliminations will yield a single candidate for the application, whose output will be unique. Insofar as other orderings have no empirical consequences, their absence from the grammar will result in an indeterminacy in the application of the rules (i.e., the equivalence class of orderings remaining after the eliminations are performed will contain more than one member), but will not result in any indeterminacy in the output. The rules are then applied (equivalently) in any one of the orders obtained, with each rule applying to the output of the previous rule.

In some of the cases discussed in the ensuing chapters where the ordering of the rules contradicts the principle of linearizability, but is consistent with the principle of local ordering just defined, we will be able to suggest the approach which a formalism for the relevant notions (especially that of "unmarked order") must take in order to allow the statement of the apparently correct solutions to the problems discussed. In other cases, however, the data available about the formal properties of such examples is insufficient to determine a unique solution to the problems of formulation, or else no particularly appealing solution presents

itself at the present. In such cases, details must be left to future research. Our aim here is not to provide a definitive theory, but rather to point out and discuss a class of problems whose correct solution must be statable in any adequate theory of grammar. The principal point of these examples is that the correct solutions are not consistent with a principle of linear order.

## Notes

<sup>1</sup> For further discussion of the issues raised here, the reader is referred to Halle (1958), Halle (1962), and the discussion in Halle (1959). This study is obviously based almost entirely on Halle's work, together with Chomsky (1967) and a few other works.

<sup>2</sup> This work, whose conclusions form the starting point for much of what follows, will frequently be referred to simply as SPE.

<sup>3</sup> The principal modification of the SPE features is the acceptance of the major class feature [<sup>+</sup> syllabic], which was described in SPE but not widely utilized there. This feature was suggested by J.C. Milner and C.J. Bailey. Other differences from the SPE features are relatively minor, and will be pointed out where they arise.

<sup>4</sup> A characterization of this theoretical orientation in 20th century American linguistics can be found, with critical discussion, in Chomsky (1964). The term "phonemic", where it is used in this study, does not in general have any reference to taxonomic phonemic theory, but rather to the redefinition of the term which has arisen in generative phonology. This re-use of a technical term with a similar, but entirely distinct meaning is perhaps unfortunate, but is clearly desirable if an unending proliferation of terminology is to be avoided. In addition, it is claimed that this usage of the term corresponds most closely to the important insights of early phonemic theorists, while avoiding the errors attendant on their precise view of language.

<sup>5</sup> Without strong constraints on the class of possible features and the kinds of rules that can relate the two levels, this claim is of course quite vacuous. For some discussion of these issues, cf. Postal (1968).

<sup>6</sup> It is assumed that the position taken in this section is no longer controversial, in light of works such as those cited in fn. 1, Chomsky (1964), and Chomsky and Halle (1967). The reader who finds them so is referred to this last-named work for a conclusive reply.

<sup>7</sup> The opposition tense/lax is employed here as a cover symbol for whatever feature may actually differentiate these two classes of vowels. From published descriptions of Telugu, it is not immediately clear what mechanism is involved. Halle has suggested (personal communication) that the relevant feature is that of [retracted tongue root], as suggested by Ladefoged (1967) for Igbo vowel harmony.

<sup>8</sup> Observe that the features characterizing this class are not identical for all of its members. This fact will prove crucial below. [č, š] are palatal-alveolars (i.e., [-ant, +cor, +high, -back, etc.]); [k<sup>y</sup>, g<sup>y</sup>, and ŋ<sup>y</sup>] are palatalized velars (i.e., [-ant, -cor, +high, -back, etc.], while [l<sup>y</sup>] is probably a palatalized dental [+ant, +cor, +high, -back, etc.]. This interpretation seems dictated by the transcription employed by Gibson.

<sup>9</sup> As noted above, the interpretation of these features is discussed in chapter 7 of SPE.

<sup>10</sup> It has been pointed out to me by Kenneth Hale (personal communication) that this formulation may not be correct. He points out that in Pame t, d, n, etc. are interdental, while c and s are alveolar. It is then likely that the n in words like /ki+ncáo/ is also alveolar, as opposed to intervocalic n (e.g., /ki+nã/). In this case, the palatalization rule will yield n<sup>(y)</sup> from n before, e.g., c, while giving ɲ<sup>y</sup> from n in isolation. It would not then be necessary for nasal assimilation to apply after palatalization in the derivation of [kinčáo]. This does not affect the point being made here, however; in this case, the palatalization rule must apply to the output of the assimilation rule, since the n only becomes alveolar through the operation of this rule. The point being made is that the representations which serve as input to one rule are created by the operation of the other. This remains the case regardless of which rule is to apply first.

<sup>11</sup> For the present, we will assume the use of the notations of abbreviation by a variable and abbreviation by parentheses and braces conform to the principles given in SPE. These principles will be discussed and evaluated in more detail in chapter IV below.

<sup>12</sup> The device of neighborhood rules was proposed by Bach (1968), and will be discussed in some detail in chapter IV below. We adapt the notation "%X\_\_Y" as an abbreviation for the sequence of two environments "/X\_\_Y" and "/Y\_\_X" where "X" and "Y" are any strings of symbols, and "X" and "Y" represent the mirror images of these strings.

<sup>13</sup> The relation "follows" is definable naturally in terms of the

relation "precedes", and will be used interchangeably with the latter in cases where it is more perspicuous.

<sup>14</sup> The reader is referred to the introductory chapters of SPE for the notions of syntactic bracketing which are relevant here.

## Chapter II - The Icelandic Vowel System

II.1 The inventory of vowels. The complete inventory of vowels in the phonetic representations of (Modern) Icelandic words of native origin consists of the following 14 elements:

1.1	<u>Phonetic form</u>	<u>Orthographic form</u>
	[iɟ]	í, ý
	[I]	i, y
	[je]	é
	[e]	e
	[aw]	á
	[a]	a
	[aɟ]	æ
	[uw]	ú
	[Ü]	u
	[ow]	ó
	[ɔ]	o
	[ö]	ö
	[ej]	ei, ey
	[öü] <sub>ː</sub>	au

This vowel system has probably had more energy expended on it in the form of diagrammatic arrangements than any other in the world. Benediktsson (1958, 1959, 1963a) has discussed the range of possible arrangements of the system and its genetic predecessors in great detail. It should be pointed out that the above represents the vowel system of stressed<sup>1</sup> syllables; in unstressed syllables, only [I], [Ü], and [a] are found, whose arrangement is uncontroversial.

A fact of the surface phonetics is the occurrence of the above stressed vowels with or without phonetic length. There has been some discussion in the literature<sup>2</sup> of the question of whether this length should



be marked distinctively on the vowel or whether it should be predicted from the adjacent consonantism: since the rule predicting the length in terms of the consonantism is needed in any case, it is possible to dispense entirely with length markings for the vowels on the basis of the following rule:

$$1.2 \quad [+ \text{syll}] \rightarrow [+ \text{long}] / \text{--- } C_o^1 \left( \begin{bmatrix} + \text{son} \\ + \text{cnt} \end{bmatrix} \right) \left\{ \begin{bmatrix} + \text{syll} \\ \# \end{bmatrix} \right\}$$

The details of this rule are fairly uncontroversial, and it will play no further role in the discussion here.

Another rule which affects the vowels inserts a glide between a vowel and certain following velars ([ŋ], [ji] from gi, [j] from gj). This rule, like that producing phonetic length, requires little comment. In addition, the phenomenon of pre-aspiration should be mentioned. Geminate stops and certain sequences of stop plus sonorant are preceded by an aspiration of sorts, with a quality similar to that of a preceding vowel. In the case of preceding diphthongs, the source of this is clear: the second element of the diphthong (the off-glide) is simply devoiced, the first element remaining unaffected. If the preceding vowel is only one mora long, however, the following occurs: the vowel segment itself is reduced to half length, and is followed by a half-length voiceless segment of the same quality. How such a process is to be formalized in linguistic theory is unclear; some device must be available to deal with the problem which is able to consider a long vowel as two morae long, and a short vowel as two half-morae, in the same rule. This process is not restricted to Icelandic; aside from the fact that the same thing takes place in some dialects of Faroese (cf. chapter III), Kenneth Hale (personal communication) has noted the exact same facts about pre-aspiration in Papago.

In examining the alternations which the vowels in 1.1 undergo, consider the forms in 1.3:

- 1.3 a. svangur 'hungry' : svengri 'hungrier'  
 b. hár 'high' : hærri 'higher'  
 c. ofan '(from) above' : efri 'higher above'  
 d. bók 'book' : bæklingur 'pamphlet'  
 e. urð 'hole' : yrðlingur 'cub'  
 f. mjúkur 'soft' : mýkri 'softer'  
 g. austur 'easterly' : eystri 'more easterly'

This alternation, which will be discussed more extensively below, is referred to as i-umlaut, and appears to be conditioned by the presence or absence of an /i/ in the syllable following the vowel affected. We can say that the vowels in the first members of each pair are basic, and that the vowels in the second members of the pairs are derived from them by an i-umlaut rule which is conditioned by the presence of a following i.

Examination quickly reveals that this statement is far from adequate to describe the facts, however. For instance, the great majority of masculine and neuter strong nouns have a dative singular which ends in -i, but most of these do not show umlaut:

- 1.4 a. hattur 'hat'; d.sg. hatti  
 b. hlátur 'laughter'; d.sg. hláti  
 c. borð 'table'; d.sg. borði  
 d. blóm 'flower'; d.sg. blómi  
 e. fundur 'discovery'; d.sg. fundi  
 f. hús 'house'; d.sg. húsi  
 g. draumur 'dream'; d.sg. draumi

The case endings are of course always unstressed; recall that the only vowels which occur in unstressed position are a, i, and u. If we assumed that the underlying vowel system in unstressed position was richer than this, and that a rule of vowel reduction applied to reduce the inventory to a, i, and u, then we could represent those phonetic unstressed i's which do not cause umlaut as some vowel other than i in the underlying representation, and then apply i-umlaut before vowel reduction. Unless some other property of these vowels becomes apparent at a later point to influence our choice, the vowel which should be chosen to represent non-umlauting unstressed i would seem logically to be one differing in as few features as possible from i. The vowel e differs from i only by the feature [high]; let us, then, represent, e.g., the dative singular ending of the nouns in 1.4 as underlying /-e/.

The umlaut rule raises still more problems, however: a number of forms appear to show umlaut, but have no i in the surface form which could have been responsible for the operation of the rule:

- 1.5    a. mann 'man (acc. sg.)'; menn (acc. pl.)  
          b. gás 'goose'; gæs 'geese'  
          c. hnót 'nut'; hnetur 'nuts'  
          d. bók 'book'; bækur 'books'  
          e. mús 'mouse'; mýs 'mice'

Confronted with these facts, we might be tempted to abandon the attempt to predict the occurrence of umlaut in phonological terms in Icelandic, and claim that it must be indicated in the lexicon by some ad hoc feature. Such an analysis has been suggested for German by Zwicky (1968) (though it has since been challenged). Though it may be the only way open, such a solution would involve an enormous cost, and some other way would certainly

be desirable if one could be found. In fact, another solution would have to lie in representing the unlauted forms in 1.5, and others where the rule must be assumed to have applied, but where there is no motivating i, with an underlying i which is later deleted. This, however, would require that every i carry some feature which indicates its status with respect to the deletion rule, since not all i's are deleted. The use of such a diacritic feature seems hardly more appealing than the use of the rule feature; however, we shall see that it can still be maintained.

To see how, let us examine the feature composition of the elements of the stressed vowel system. Considering only the vowels á, é, í, ó, ú (and ý=f), and a, e, i, o, u (and y=i), we see that the latter are composed, basically, of the former minus a glide.<sup>3</sup> Though there are virtually no alternations in the language between members of the one set and those of the other, an economy of representation could be obtained if we eliminated the underlying diphthongs, replacing them with representations identical to those of the corresponding simple vowels except for the presence of a feature to indicate that they undergo a rule inserting a glide. Since there is, in addition to the matter of the presence of the glide, a phonetic difference in tenseness between the two sets, we can use the feature [ $\pm$  tense] to indicate susceptibility to the rule of diphthongization.

There are some few generalizations that can be made about the distribution of the feature [ $\pm$  tense] in underlying forms,<sup>4</sup> but its value in syllables that will receive the stress is generally unpredictable. In unstressed syllables, however, only [- tense] vowels appear, suggesting the existence of a rule like 1.6 as part of the vowel reduction process:

$$1.6 \quad [+ \text{syll}] \rightarrow [- \text{tense}] / \left[ \begin{array}{c} \text{---} \\ - \text{stress} \end{array} \right]$$

Note that we now have available to us a feature which can be used to characterize the difference between those (unstressed) i's which, after causing umlaut, are deleted, and those which remain. If the i's which are to be subject to deletion are represented as [- tense], while those that are not are [+ tense], we can formulate a rule which deleted unstressed lax i, and apply it before 1.6.

This rule deleting unstressed lax i (and u as well, as will be suggested below) does not apply in all cases, but only in certain phonologically formulable environments. In other environments, the distinction of tenseness in unstressed vowels is irrelevant, and is neutralized by the operation of rule 1.6. The exact formulation of the vowel deletion rule will have to await further discussion of the umlaut rules, since it is their operation that provides the crucial evidence for the presence of the deleted vowels. At this point, however, we can say that the vowel inventory contains both tense and lax a, e, i, o, and u, in addition to [ö], [ej], and [öy], which have not yet been discussed. It has further been established that the unstressed vowel system is not restricted to a, i, and u in underlying forms, but contains in addition at least some tense vowels, plus the vowel e.

II.2 The u-umlaut rule. One of the most interesting and productive alternations in the Icelandic vowel system is that between a and ö in forms like these:

- 2.1 a. fata 'pail'; oblique singular cases fötu
- b. heilagur 'holy'; nom. sg. fem. heilög

As i-umlaut was conditioned by the presence in the following syllable of an i, so this alternation appears to be produced by a following u. But this

formulation leaves a number of facts to be accounted for: first, a great many nouns end in -ur in the nominative singular but do not show u-umlaut:

- 2.2    a. dalur 'valley'  
          b. akur 'field'  
          c. faraldur 'epidemic'

These words do show u-umlaut when followed by other endings containing u, such as the um of the dative plural; it is thus definitely a characteristic of the ending ur and not of the root that the forms in 2.2 do not show umlaut.

Along the lines of the proposal advanced above for dealing with the dative singular as /-e/, we might propose to represent the vowel of -ur as something other than u (e.g., o), which does not otherwise occur in unstressed position, and then have vowel reduction convert this other vowel to u after the u-umlaut rule has applied. This solution will lead to some loss of generality in dealing with the ending -ur, however. After nouns which end in a vowel, this ending has the shape -r, with no vowel:

- 2.3    a. læk<sup>n</sup>ir (from /lákni+r/) 'doctor'  
          b. snjór (from /sneo+r/) 'snow'

If the nom. sg. ending is represented as, say, /-or/, a rule will be needed to delete the vowel after a vowel stem. Other rules are necessary to delete vowels before vowels, but not after. In addition to this fact, the ending which appears as -ur after consonants and as simply -r after vowels is completely assimilated after certain dentals (r, l, n, or s). Thus, the nom. sg. ending appears as  $\emptyset$  after r or s (by an independently needed rule which converts rr to r and ss to s), as l after l, and as n after n:

- 2.4 a. húmar (from /humar+r/) 'lobster'  
 b. djöfull (from /deful+r/) 'devil'  
 c. jötunn (from /etun+r/) 'giant'  
 d. ís (from /ís+r/) 'ice'  
 e. lax ([laks]) (from /laks+r/) 'salmon'

Thus, the vowel deletion rule would also have to apply after these segments, in order to reduce /-or/ to simply /-r/. We could also, however, derive all of these instances of -ur, etc., from an underlying /-r/, and incorporate in the grammar the rule 2.5 of u-epenthesis:

$$2.5 \quad \emptyset \rightarrow u / C \_\_ r \left\{ \begin{array}{l} C \\ \# \end{array} \right\}$$

This rule applies after all of the assimilations just mentioned (which account for, e.g., 2.4) have applied. It is, in fact, a general property of Icelandic words that Cr sequences do not occur before either C or #.

As was the case with the effects of the i-umlaut rule, there are Icelandic forms which appear to show the effects of the u-umlaut rule, without there being present in the phonetic representation any other reflex of a u that might have been responsible:

- 2.6 a. barn 'child'; börn 'children'  
 b. blað 'newspaper'; blöð 'newspapers'  
 c. tagl 'horsetail'; tögl 'horsetails'

A solution to this problem along the lines of that suggested for the analogous problem with respect to the operation of the i-umlaut rule seems indicated. That is, u in unstressed position may be considered to be either [+ tense]; then in certain positions, [- tense] u will be deleted, presumably by the same rule that is responsible for the deletion of [- tense] i, while [+ tense] u will be made [- tense] by rule 1.6.

The u-umlaut rule produces not only a surface a/ö alternation, but also an alternation between a and u. This latter alternation is confined to unstressed position:

- 2.7 a. meðal 'drug, medicine'; dat. pl. meðulum  
 b. hérað 'region'; dat. pl. héruðum

In addition, more than one vowel may be affected by the rule; but only if no stressed vowel or vowel other than underlying a intervenes between the vowel affected and the u producing the change:

- 2.8 a. fatnað 'suit of clothes'; dat. pl. fötnuðum  
 b. almanak 'calendar'; dat. pl. almanökum<sup>5</sup>

There appear to be only two exceptions to this rule:

- 2.9 a. kafald 'blizzard'; nom. pl. köföld  
 b. hafald 'lifting gear, jack'; nom. pl. höföld

These can presumably be handled by the same exception mechanisms as were suggested in fn. 5 to account for some of the possible forms of the dative plural of banani.

Notice that in light of what has been suggested so far about the vowel system of unstressed syllables (i.e., that it is richer than just the three elements a, i, and u, and that there exists a rule to reduce other vowels to one of these), an alternation between a and u is just the reflex we would expect of an alternation between a and ö in unstressed position. If we permit the u-umlaut rule to send a to ö in unstressed, as well as stressed position, we can have the vowel reduction rule convert the ö to u. On this basis, we could formulate the vowel reduction rule approximately as follows:



$$2.10 \quad [+ \text{syll}] \rightarrow \begin{bmatrix} + \text{ high} \\ - \text{ tense} \\ \alpha \text{ back} \end{bmatrix} / \begin{bmatrix} \text{---} \\ \alpha \text{ round} \\ - \text{ stress} \\ - \text{ low} \end{bmatrix}$$

Note that this rule incorporates the effects of 1.6; there is apparently no reason to assume a tenseness distinction in unstressed a, since such a distinction would have no effect, so the effect of 2.10 is to assimilate backness to rounding in non-low unstressed vowels, and to make these high and lax as well.

As was noted above, if a vowel other than a intervenes between an a and a following u, the a is not affected:

2.11 akkeri 'anchor'; dat. pl. akkerum

The facts which have been adduced so far concerning the rule of u-umlaut suggest that it can be formulated approximately as in 2.12:

$$2.12 \quad [+ \text{syll}] \rightarrow \begin{bmatrix} + \text{ round} \\ - \text{ back} \\ - \text{ low} \end{bmatrix} / \begin{bmatrix} \text{---} \\ - \text{ tense} \\ + \text{ low} \end{bmatrix} (C_o \begin{bmatrix} + \text{ syll} \\ - \text{ stress} \\ + \text{ low} \end{bmatrix})_o C_o \begin{bmatrix} + \text{ syll} \\ + \text{ high} \\ + \text{ back} \end{bmatrix}$$

The portion of the rule in parentheses (which represents  $(C_o \underline{a})_o$ ) represents the abbreviation of a potentially infinite schema of rules, as suggested in SPE. It will be suggested below in chapter IV that at least some uses of this device should be replaced with iteratively applied rules, but this does not appear to be such a case. There do not exist cases which would conclusively require the full power of the infinite schema or anything near it; as nearly as can be determined, there are no forms in the language in which more than two vowels are affected. One might then propose that the 'infinite schema' device be replaced by a single instance of  $C_o \underline{a}$  in parentheses; that is, that 2.12 contain precisely two rules. If this approach

were taken, however, the two rules would be formally related by a notation which should enforce disjunctive order on them (see chapter IV for a justification of this claim). In fact, native speakers of Icelandic, confronted with nonsense words which contain three consecutive syllables with a, generally agree that more than two vowels can be affected. This implies that their internalized grammars contain a rule formally like 2.12, rather than a rule schema representing only two rules, as has just been suggested. Since this latter would seem to be the simpler solution, were it not for the principle of disjunctive order which prevents it, the fact that it does not seem to have been adopted constitutes a strong sort of evidence for the psychological reality of this principle.

Granted the formulation of the structural description of the u-umlaut rule, we can inquire into the possibility of a simplification in its structural change. As a statement of the facts of the surface phonetic alternation it is perfectly adequate, but the change from a to ö is sufficiently unusual to suggest that another formulation might be preferable. The rule actually performs three changes: the affected vowel becomes [+ round], [- low], and [- back]. Let us consider each part of this change.

First we note the absence of any other segments with the features [+ low, + round]. If the umlaut rule performed only the change to [+ round], its outputs would be unambiguously marked, and some later rule or rules converting these vowels into the phonetic [ö] would not need to affect any other segments. In fact, there is some evidence to suggest that the vowel ö is not a front vowel at every stage of its history in derivations: the rule of velar fronting, which will be mentioned below, applies generally before front vowels regardless of their source, but does not apply before ö or [Ü]. This leads us to suggest that these vowels should not be made front until

after the velar fronting rule.

We can also note, as a late phonetic fact, that the vowel represented by underlying o is pronounced as a low vowel [ɔ]. The grammar will thus have to contain at some point a rule providing this specification. If we assume that the vowels produced by the umlaut rule are, immediately after the operation of the rule, [+low, +round, +back], the rule which makes o into a low vowel can be formulated as a vowel shift rule, which interchanges the two non-high round back vowels:

$$2.13 \quad [+ \text{round}] \rightarrow [-\alpha \text{ low}] / \left[ \begin{array}{c} \text{---} \\ + \text{syll} \\ - \text{high} \\ - \text{tense} \\ \alpha \text{ low} \end{array} \right]$$

If we assume that the surface vowel ø starts its life as [ɔ], then becomes [o] by 2.13, and then is fronted to [ø] by some later rule, we could considerably simplify the rule of vowel reduction (2.10) by allowing it to apply at the point just after 2.13: at this point, there are no [-low] vowels that disagree in backness and rounding, so it would be unnecessary to assimilate these features:

$$2.14 \quad [+ \text{syll}] \rightarrow \left[ \begin{array}{c} + \text{high} \\ - \text{tense} \end{array} \right] / \left[ \begin{array}{c} \text{---} \\ - \text{stress} \\ - \text{low} \end{array} \right]$$

The facts just presented indicate that the vowel shift rule just formulated (2.13) has some justification in a synchronic grammar of Modern Icelandic; there is reason to believe, however, that it has been part of the language for some time. Aside from the facts which will be presented below concerning the rule of breaking, there are several instances of the vowel ó whose etymology is somewhat problematic. For example, consider the following

Old Icel. forms:

- 2.15 a. nótt 'night' from *nǫgt*, from *nahtu*  
 b. Óttarr 'pers. name' from *ǫhtuharir*  
 c. ó 'river' (alt. form of *ǫ*) from *ahwu*

The development of these forms is as follows: u-umlaut (a rule which was operative in Old Icel., as well as in Mod. Icel., and which had the effect simply of rounding an *a*, as we assume here is the correct formulation of the modern rule) affected the initial *a*, converting it to *ǫ* ([ɔ]); this vowel was then lengthened by a rule applying before clusters of h+consonant, and then somehow the resultant *ǫ* was raised to *ó*. The normal development of long *ǫ* is *á*, with the result that the u-umlaut rule came to have no effect on long vowels (as it had had originally), and in Mod. Icel. is restricted in application to short vowels. Grammars such as Heusler (1964) attribute this raising to the influence of "ein benachbarter Nasal," and this is a conceivable explanation of some of the forms. It is not a possible explanation of, e.g., 2.15c, however, where there is no nasal present. This seems to imply that there was already some relation between low and non-low vowels in the non-high back rounded set at an early point. If we assume that a rule like 2.13 was applicable in Old Icel. as well, we can construct the following explanation for 2.15: in these forms, the vowel lengthening applied after, rather than before the vowel shift in these forms. The vowel shift is normally restricted to the short vowels, and normally applies only after all lengthening processes. If in these forms, however, it were allowed (exceptionally) to apply to the intermediate stage, before the lengthening before h+C, this would have the effect of lengthening the derived vowel ([ɔ]), yielding *ó* rather than the expected *ǫ*.

An additional instance in support of this suggestion is the

problematic fact, observed by Sturtevant (1953), that there were two suffixes óttr in Old Icel., with different properties. One of them, a native Germanic suffix, and presumably not analyzable, had no effect on the preceding word; another suffix, however, with the exact same phonetic shape, had the effect of producing u-umlaut in a preceding vowel a. This second suffix was the reflex of the borrowed Latin suffix -actus. In Germanic, the c of -actus would have become h; the final s would have become R. At a later point the cluster -ht- normally became -tt- (these last two changes are Scandinavian, rather than Germanic, of course). If we assume the suffix still had the representation /-ahtuR/ at the relevant point in Old Icel., this would of course produce u-umlaut in a preceding a; it would also produce u-umlaut in the a of /-ahtuR/. After this, we could assume that this suffix, like the forms in 2.15, underwent vowel shift exceptionally early, yielding /-ohtuR/; then the lengthening rule would apply, followed by the rule of -ht- → -tt-, followed by the loss of unstressed lax u, followed by the reduction of R to r. Thus, the output from underlying /-ahtuR/ would be exactly an umlauting suffix with the phonetic shape -óttr. This seems to be the only principled explanation of the suffix in question; furthermore it is impossible to invoke any principle involving nasals to account for the quality of the vowel. Thus, this suffix, like the forms in 2.15, seems to indicate that even in Old Icel. there was a vowel shift rule similar to 2.13 whose application at exceptional points in the sequence of rules was responsible for a number of puzzling exceptional forms.

Accepting rule 2.13, then, and having it apply to the output of the u-umlaut rule, the [- low] quality of the output vowel [ö] is accounted for. It still remains to account for the fact that this vowel is [- back] as well. Thus, we will need a rule like 2.16 as well:

2.16

$$[+ \text{syll}] \rightarrow [- \text{back}] / \left[ \begin{array}{c} \text{---} \\ - \text{high} \\ - \text{low} \\ - \text{tense} \end{array} \right]$$

Since the phonetic value of the umlauted vowel ö has been [ö] for some time, the rule 2.16 must also have been part of the grammar since relatively early times (cf. Benediktsson, 1958, and others for details of the history). But in this case, we can explain the later fronting of lax u to [ʊ] as a generalization of 2.16 to 2.17.

$$2.17 \quad [+ \text{syll}] \rightarrow [- \text{back}] / \left[ \begin{array}{c} \text{---} \\ - \text{low} \\ - \text{tense} \end{array} \right]$$

which is the present form of the rule. It is the fact that this rule is independently necessary to explain the phonetic value of u that permits the vowel shift rule to be maintained in its present form.

We have now justified, via rules 2.13 and 2.17, the elimination of the specifications [- low] and [- back] from the structural change of rule 2.12 (u-umlaut). This rule now looks like this:

2.18

$$[+ \text{syll}] \rightarrow [+ \text{round}] / \left[ \begin{array}{c} \text{---} \\ + \text{low} \\ - \text{tense} \end{array} \right] (C_o \left[ \begin{array}{c} + \text{syll} \\ + \text{low} \\ - \text{stress} \end{array} \right])_o C_o \left[ \begin{array}{c} + \text{syll} \\ + \text{high} \\ + \text{back} \end{array} \right]$$

We have succeeded in explaining the remainder of the change in terms of otherwise needed rules of the grammar.

II.3 The order of u-umlaut, and Old Norse Skaldic verse. The place of the u-umlaut rule in the ordering of Icelandic rules presents an interesting situation. There exists a rule in Icelandic (whose precise formulation will not be attempted here) whose effect is to delete a vowel between certain consonants and a following sonorant consonant before endings beginning with

a vowel. This rule is responsible for the alternations in 3.1:

- 3.1 a. hamar 'hammer'; dat. sg. hamri  
 b. fífill 'dandelion'; dat. sg. fífli  
 c. gimbur 'ewe'; gen. sg. gimbrar

When we examine forms such as those in 3.2, we see that the syncope rule must evidently apply before the u-umlaut rule, for until the syncope applies, a vowel intervenes between the a of the root and the u of the ending which would prevent the rule from applying. It is only the deletion of this vowel by syncope that enables u-umlaut to apply:

- 3.2 a. alin 'ell (measure of cloth)'; dat. pl. ölnu<sup>6</sup>  
 b. ketill 'kettle' (dat. sg. katli); dat. pl. kötlum

Since u-umlaut does apply in these forms, we must conclude that it does so after the application of syncope.

But now let us consider another class of forms, composed generally of roots plus one of the suffixes /-ul/, /-ur/, or /-un/:

- 3.3 a. böggul 'package' (cf. baggi 'bag'); dat. sg. böggli  
 b. djöfull 'devil' (from /deful+r/, = /djaful+r/;  
 dat. sg. djöfli  
 c. jötunn 'giant' (cf. eta 'to eat' (n.b. giants eat people));  
 dat. sg. jötni

In these cases the u of the suffix produces umlaut, even though it is itself deleted by the operation of the syncope rule. If this is indeed the case, then it cannot be that the syncope rule precedes the u-umlaut rule (as was found necessary for the forms of 3.2) in all forms.

One approach to the ordering paradox just discovered would be to deny that both of the rules in question (syncope and umlaut) apply in the forms in 3.2 or the forms in 3.3. It seems obvious that syncope applies in both

classes; it is also fairly clear that umlaut must apply in the forms in 3.2. It might be possible, however, to deny that umlaut ever applies to the forms in 3.3: since these words will have the vowel ö in all of their forms, one might argue, along the lines of Kiparsky (1968a), that they should contain this vowel in their underlying representations as well. While this would run counter to the clear intuition that they are directly related to other forms of the same root without umlauted vowels, and hence should be represented in the same way (that is, without umlauted vowels), there is stronger evidence than this against such a proposal. To see this, we must consider the system of Skaldic poetry practised for a great many years in the Old Norse area. While this does not, of course, provide evidence about the modern language, it can easily be shown that the older form of the language contained a u-umlaut rule identical in all relevant respects with the modern one, and that the same problems with respect to its ordering relative to the syncope rule existed in Old Norse as in the modern language. Thus, if it can be shown that both umlaut and syncope applied in Old Norse forms parallel to 3.3 as well as those parallel to 3.2, this will constitute strong evidence for the same analysis of modern Icelandic forms, which is confirmed by the desirability of maintaining a uniform representation for the roots of 3.3.

The metrical tradition of Skaldic verse was undoubtedly developed as an elaboration on the earlier, freer meters of early Germanic and the verse of the Eddic poems. Though there do appear a few lines in early runic material which may be Skaldic in form, the style appears to emerge nearly full-blown in the work of Bragi hinn gamli, a Norwegian of the middle 800's. This style was variously elaborated, especially by Egill Skallagrímsson (immortalized in Egils Saga Skallagrímssonar) ca. 910-990, and can be regarded as fully matured in the verse of Sighvatr Þórðarson, ca. 995-1045.



As a medium for the composition of occasional praise, heroic description, magical incantations, love poetry, etc., it flourished for the next several centuries. In ca. 1223, Snorri Sturluson felt that the art of versifying was being practised by less skilled hands than previously, and wrote a long treatise on poetics (the hattatál, or 'account of meters', forming part of Snorri's prose Edda), which was designed to enlighten his contemporaries as to the correct bases of Norse poetics. This is still one of the most valuable studies of Skaldic verse form in existence, and is fundamental to all later descriptions. More will be said below of the influence of this work.

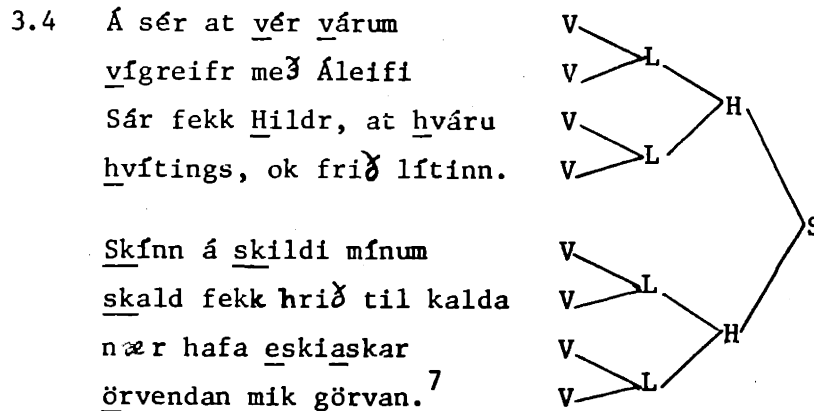
Since Skaldic verse was primarily practised by professional poets, whose livelihood depended in many instances on their ability to compose a well turned stanza or two in celebration of a lord's deeds or whatnot, it quickly became highly intricate and conventionalized, especially in its syntax. The sentences making up a stanza could be interwoven in fantastically complex ways; to date, very little of any significance has been said in the line of a Skaldic syntax, and such a task will hardly be attempted in this excursus.

The basic meter of Skaldic verse was known as drottkvætt, or court poetry. The basic unit of this form was the stanza or vísa. A vísa, in turn, was composed of two metrically identical units each called a helming, or half stanza; a helming contained two couplets, the long lines of other Germanic verse. The two long lines of a helming are in turn metrically similar in most cases. Each of them is composed of two half lines, called a vísuorð, which is the basic unit in terms of which the poetry is constructed. As will be seen, each of these units has a certain amount of ontological respectability, since for each there are metrical constraints whose statement is crucially in terms of such a unit.

Leaving aside the half line or *vísuorð* for the moment, we see that one of the most important aspects of the poetry is defined over the unit of the long line -- the initial alliteration, which is the main tie between Skaldic and other Germanic verse. The basis of the alliteration is the requirement that within a metrical long line, at least three stressed syllables must begin with the same consonantism. One of these must be the first syllable of the second half fine of the long line (the so-called *höfuðstafr*); the other two (called *stuðlar*, or 'supports') must come in the first half line. This alliterative identity applies between the initial segments of the syllables in question, if these are [+ consonantal]; if, furthermore, this segment is s, and it is followed by a stop, the two stops must also be identical (i.e., sp, st, and sk rhyme only with themselves). Words whose initial segment is [- consonantal] (i.e., vowels and glides) alliterate with one another regardless of their other features. To account for this latter fact, Jakobson (1963) postulates a 'zero phoneme' (a glide which is lax and has no other features) which is said to precede such segments. Traditionally, it has been assumed that the glottal stop which characterizes onsets in German was general in those Germanic languages where such a rule is found in the poetry, and that it is this glottal stop which alliterates, reducing the so-called vowel alliteration to a special case of ordinary consonantal alliteration. Jakobson points out, however, that there is no evidence for this other than the poetry, and in the only language where such poetry is still composed according to such rules (modern Icelandic) it is phonetically absent. Jakobson's alternative of the 'zero phoneme' has the merit of being phonetically unobservable, and hence not disconfirmable by any possible empirical evidence, but again there is no evidence for these mythical segments save the possibility of simplifying the alliteration rule for the

poetry, the phenomenon its presence is supposed to account for. Kiparsky has analyzed (1968c) the constraint as a required identity of analysis under the schema  $\#C_0^1$  (where this must be further complicated to take into account the s+stop cases). This provides a unitary statement of the constraint without the metaphysics of 'lax featureless onglides'.

As an example of a Skaldic stanza, consider the following:



--þormóðr Kólbrunaskáld

In this example, the underlined segments carry the alliteration; the units labelled V are vísuorð, or half lines; the units L, long lines; H, helmingar; and the S is a stanza.

The structure of the half line, and especially the distribution of stress and length over it, has been studied in considerable detail. The structural tendency is clearly toward a trochaic line of three feet, the downbeat of each foot being a long syllable. There are a large number of exceptions to this pattern however; Sievers (1878,9) has catalogued most of them, and has attempted to present rules for the composition of the line. This treatment, while compendious and probably complete, seems to me to miss several essential points: the pattern of the verse as imposed by strictly metrical considerations is basically rather simple, and the complexities of

the pattern found by Sievers are due rather to the consequences of the requirements of alliteration and rhyme (to be discussed below), in conjunction with the structure of Old Norse. Though it is well beyond the scope of this excursus to document this claim, the abstract pattern of the verse<sup>8</sup> appears to be the following:

### 3.5    w w    : w S

In this formula, w is realized in general by a single syllable, with the following exceptions: a) the first w may be realized optionally by two short syllables not separated by #, or by a short syllable followed by an unstressed monosyllable; b) any w may optionally be realized by two of three consecutive vowels. The S in the formula is realized by two syllables at the end of a word, of which the first is long and the second does not bear stronger stress than the first. Because of the structural properties of the language, this will almost always mean that the second is short and unstressed. In addition to the above, there is an added constraint that there be at least two instances of main (i.e., primary or secondary) stress within the scope of the four w's.

On this pattern are superimposed added constraints of a rather restrictive type. In all such poetry, the alliteration described above is compulsory; in addition, in regular drottkvætt, there are requirements of internal rhyme as well. In the first half line of each long line there must be a stressed syllable among the w's with the same final consonantism as the first syllable of the S, but with different vocalism; in the second half line there must be such consonant rhyme in syllables with the same vocalism. In the Snorra Edda there are illustrated numerous further possible complexities: further internal rhyme requirements, end rhyme (runhending)

instead of internal rhyme, etc. One can combine any of these devices, sometimes with strengthened alliteration requirements, and Snorri gives examples of many different variations, some so complex and seemingly restrictive as to preclude their use for anything but occasional novelty doggerel.

The basic patterns of Skaldic verse, then, can be described as variations on the pattern 3.5, (involving occasionally the addition of another w or an S\* (— a single stressed long syllable between the two groups of w's in 3.5) and/or variations on the realization rule for S (e.g., it may be made the same as that for S\*)). This basic pattern is similar to that of the Latvian trochaic folk songs and some traditional Serbo-Croatian poetry (cf. Zeps, 1963; Jakobson, 1952; and Halle, 1967, for discussion), with the exception that the S\* is the same as the S in these forms. While these patterns are of course very general, simple arrangements of units, the fact is that other Indo-European poetic traditions, such as Sanskrit, Greek and Latin poetry, have used considerably different ones; furthermore, the mapping rule for S is duplicated in Scandinavian, Slavic and Baltic verse. These facts perhaps indicate some interesting community of poetic tradition, including Slavic, certain Baltic languages, and the earliest distinctively Scandinavian verse: such a community would not be unlikely in light of the extensive explorations of the Norsemen in Russia and the Baltic regions at approximately the time that would be necessary to establish the presumed community.

The fact that Skaldic verse was in general very carefully composed makes it possible to use it for evidence for certain aspects of the structure of the language. In particular, it will be remembered that the rules of internal rhyme require identity between certain vowels of the even half lines.

If a line does not exhibit such an identity, apparently, it may simply be that the poet has composed a metrically deviant line; such a thing is of course perfectly possible when the restrictions on correct lines are as narrow as they are in Skaldic stanzas. If, however, the deviant lines fall into reasonably well defined classes, then we must assume that they constitute a special sort of regularity, and we should attempt to relate such regularities to the structure of the language in general. An example of such reasoning is found in Kiparsky (1968c), where it is shown that certain apparent deviations in the lines of the Kalevala can be explained if we assume that the meter is defined on a representation to which certain of the rules of the phonology have applied, but to which other rules have not yet applied. This representation is neither phonemic nor phonetic, but an intermediate level with no other known systematic properties; in particular, it is not identifiable with the level of taxonomic phonemics. A similar argument is found in Zeps (1963).

A very large class of exceptions to the rules of internal rhyme consists of lines in which one of the rhyming syllables has the vowel a and the other has ö (= o, not ø which later merged with o). In a representative corpus of approximately 4,000 lines of Skaldic verse (Jónsson, 1929) approximately 94% of the 'unmetrical' lines were of this type. An example of such a 'deviant' line is:

3.6 allvaldr, nai<sup>ø</sup> sköldum

This type of line is very common, and constitutes a well defined class of exceptions. If the rule of regular rhyme is to be preserved, we must find some level of phonological representation at which the vowels are identical. It is, in fact, true that all instances of ö that rhyme with a can be derived from a by the u-umlaut rule. Let us assume, then, that the metrically

relevant level of representation is prior to the application of this rule. This level cannot be identical with that of underlying representation, as is shown by rhymes like

3.7 skjöldum : aldri

for the vowel in the deepest level of representation of skjöldum is e, as shown by its other forms (e.g., dat. sg. skildi, where underlying e is raised to i before the exceptional dative ending /-i/). This e is broken to ea, which becomes ja, which is then affected by the umlaut rule. (These processes will be described below in section II.5 under the heading of breaking.) It would seem, then, that the metrically relevant level is before the application of u-umlaut, but after the various parts of breaking.

This explanation appears to be quite satisfactory, if all instances of rhyming ö are derived from a. All will then appear as a at some point. In Kiparsky (1968a) however, a proposal was made which would make such a solution impossible if adhered to. In discussing a naturalness convention for phonological representations, Kiparsky suggests the constraint that if a form shows some phonological feature in all of its variants, that feature must be assumed to be present in its underlying representation as well. As applied to this case, this would require that for any word where the vowel is ö in all of the forms, regardless of whether it would be possible to derive these instances of ö from underlying a, the underlying vowel must be ö. This does not affect the forms adduced so far; all of them show vocalic alternations which require the positing of underlying a. There are, however, rhymes such as:

3.8 a. knörr : varrar

b. juföl : dvöldu

c. örleiks : jörlum

In the first member of each of these pairs, it is possible, but not necessary, to represent them with underlying a vocalism. On Kiparsky's proposal, then, they must have underlying ö, and never undergo a stage where they have a. The second member of each pair is one where the underlying vowel must be a. While these unproductive instances of ö are not too common in rhymes, they are about as common as their declension types are in the total distribution of forms in the language. If we assume that they have non-derived, underlying ö, then they will not have the same representation as the second members of the rhymes in 3.8 at the point defined above: namely, after breaking but before u-umlaut. Rhymes such as 3.8a, in fact, will never have identical vocalism, at any point of the derivation. We seem to have two possible alternatives open to us: either we reject Kiparsky's proposal, and allow all instances of rhyming ö to be derived from a, preserving the unity of the statement about rhyme; or we maintain Kiparsky's proposal, and add an ad hoc rider to the statement of the rhyme constraint to the effect that while other vowels require identity as to rhyme, the vowels a and ö rhyme with one another. Such a solution as this second one may in fact be forced on us, but it would miss the fact that precisely the two vowels related as exceptions are also related in a non-arbitrary way, by a highly productive device of the phonology of the language.

There are, however, additional facts that show that such a solution, whether desirable or not, is inadequate, and that only a rejection of Kiparsky's proposal, at least in its full generality, will permit the constraints on internal rhyme to be stated at all. As was noted by Kahle (1892), there are two sources for the vowel ø in Old Norse: most commonly, it arose as the i-umlaut of underlying o, but it could also come from rounding of underlying e by a following v (which comes from w, from underlying



u). Forms with the former source for ø, such as før, rhyme freely with words of the latter type, such as the forms of the verb gøra, but an additional peculiarity of these latter is that they can also rhyme with words with the vowel e, such as sverð. Krahe takes this simply as an indication of the date of the rounding of the e, and claims that the vowel of gøra was not yet rounded at the point such poetry was written. This, however, is plainly inconsistent with the fact that during the same period gøra rhymes with, e.g., før, which clearly had a rounded vowel from underlying o. The only way to state these facts is to say that the rounding rule (which is clearly very closely related to the u-umlaut rule) may create new rhymes, but that forms that rhyme before its operation are also metrically correct. In this case, it would be impossible to deal with the irregular rhymes of ø from e with e by an ad hoc rider to the effect that ø rhymes with e, because only certain instances of ø do: namely, those that are derived from an underlying e. Since all instances of ö that rhymed with a have (potentially) the same source, it was not possible to give such an argument directly for that vowel.

It appears, then, that the only way to state the rhyme rule for Skaldic poetry is approximately as follows:

- 3.9 Two syllables rhyme if they have the same final consonantism, and if they have the same vowel before or after the operation of the rounding/u-umlaut rule.

If this is so, we must restrict the application of Kiparsky's naturalness convention in such a way as to permit all rhyming ö to be derived from a. There are various proposals one might make to accomplish this; it is sufficient to note here that this implies that all such words involve the rule of u-umlaut in their derivation. It will be remembered that the point of this entire

discussion has been to show that certain instances of the vowel ö which Kiparsky's convention would require to be represented as underived (namely, the forms in 3.3) are in fact derived from underlying a. Since the forms in question appear in Skaldic rhymes with a and with derived ö, we have accomplished this.

To conclude this discussion of Old Norse poetics, a few historical observations are in order. Benediktsson (1960) also noted the facts adduced above about the rhymes of a and ö (as he has noted so many of the most interesting facts of the language), and attempted to give an explanation of them which would avoid the purely ad hoc statement "a rhymes with ö, too." His explanation was in terms of their taxonomic phonemic status: he claimed that while the other (taxonomic) phonemes of Old Norse were completely opposed to one another, the opposition between /a/ and /ö/ is neutralized in a very common position: before a u in the following syllable. This is due to the operation of the u-umlaut rule, of course. Because of this widespread neutralization of the taxonomic opposition, Benediktsson claims, these two phonemes were most likely to be identified among the units of the language.

Benediktsson further notes the following: in the work of Snorri Sturluson and later poets, there are no more instances of rhymes of ö and a than are explicable as purely errors. It is further the case that the neutralization of the opposition between /a/ and /ö/ before a following u was lost when the rule of u-epenthesis (rule 2.5 above) was introduced into the language: since this kind of u did not cause u-umlaut, the taxonomic phonemic opposition of /a/ and /ö/ before /u/ was restored. If this event were correlated with the loss of a/ö rhymes, Benediktsson would have found a very important piece of evidence for his theory of the nature of such rhymes.

Unfortunately, however, the two events are not simultaneous. As Benediktsson observes, both are susceptible of a reasonably precise dating, and there is a period of about 75 years between the loss of the odd rhymes and the introduction of the new u's. Benediktsson has no satisfactory explanation for this fact, and it seems to be a crucial objection to his theory. Another explanation for the loss of the a/ö rhymes in poetry of Snorri Sturluson and later poets is available, however: namely, the influence of Snorri's work itself. Snorri was concerned to produce a manual for poets that could be followed more or less mechanically to turn out correct stanzas. As such, he was concerned (in characterizing the rules of the poetry) with giving a set of generalizations about the surface forms of such poetry. By becoming reflective about the surface generalizations to be extracted from it, he lost sight of the more abstract components of the grammar of the language that had guided earlier poets, whose skill was largely unconscious. As a good Norse linguist, in any event, he was probably inclined to the taxonomic; while he cannot be identified with the so-called 'first grammarian', whose work was undoubtedly the earliest and clearest example of a straightforward taxonomic phonemic analysis on record (cf. Haugen, 1950), he was no doubt influenced by the same attitudes. As such, it was precisely the codification of the rules for writing Skaldic verse that resulted in the elimination of one of its features: since Snorri could not see how a/ö rhymes could be considered regular, he ruled them out, and since his work was the basis of all poetizing from then on, they perforce disappeared from the verse.

We have, then, established that in the derivation of forms like those of 3.3, the u-umlaut rule does apply; since the syncope rule also applies, we must resolve the paradox that was noted at first. This could of course

be done by building the environment of the syncope rule into the structural description of the u-umlaut rule, so that a vowel which was to be syncopated would be skipped over by that rule. This revised u-umlaut rule could then apply uniquely before the syncope rule, giving both 3.2 and 3.3. This environment is by no means a simple one, however, and to build it into the u-umlaut rule, in addition to its own appearance in the grammar, would be tremendously expensive; indeed, it should be, for it misses a tremendous generalization.

The correct relation between these two rules seems to hinge on the different characteristics of the two sets of forms 3.2 and 3.3. In the forms in 3.2, the order syncope - u-umlaut is a feeding order, since the syncope rule turns these forms into the correct shape for the application of the u-umlaut rule. Thus, for these forms, the order syncope - u-umlaut is an unmarked order. For the forms in 3.3, on the other hand, the order syncope - u-umlaut would result in the forms being removed from the class of those to which u-umlaut could apply, by the syncopation of the u which should cause umlaut. For these forms, then, the order which is a feeding order for the forms in 3.2 is a bleeding order; and, hence, a marked order. If we operate on the theoretical assumptions about order which were outlined in chapter I, and simply do not make any statement about the order of the two rules in the grammar of Icelandic, then they will apply in each case in the unmarked order. This order will be different for the two classes of forms.

In the course of this discussion, we have revealed two facts about the notion of 'unmarked order': first, that the notion cannot be defined between a pair of rules in isolation, but is (at least) a ternary relation, defined on the basis of two rules and a form to which they are to apply.

That is, the shape of the particular form may influence which of the possible orderings is the unmarked one. Secondly, as a consequence of this, we see that there is an empirical difference between the theory assumed here (which permits unspecified orderings to be taken as unmarked) and one which would require a specification of the relative ordering of every pair of rules, and then apply interpretive principles to evaluate this order as marked or unmarked. This difference resides in the fact that no single ordering may be evaluable as unambiguously marked or unmarked, if the possible shapes of forms can influence the question.

The traditional explanation of the process of u-umlaut (for which see Noreen, 1923; Gutenbrunner, 1951; Heusler, 1964) claims that umlaut of unstressed a to u took place first, and that these secondary u's then caused the umlaut of a preceding stressed a to ö. The problems with this solution are at least the following: first, it requires two separate and unrelated rules to account for the changes; secondly, it makes the claim that the change in stressed syllables is only accidentally related to the change in unstressed syllables. There is no reason why we could not have had, say, a becomes i in unstressed position followed by the same u-umlaut rule in stressed position. The solution suggested here preserves the unitary effect of the change, and accounts for the surface dissimilarity of effect in stressed and unstressed positions in terms of the otherwise motivated rule of vowel reduction. Thirdly, the traditional solution would make it impossible to account for the exceptional forms like 2.9, where vowel reduction does not apply. Instead of being simple exceptions, these would require a special minor rule. The truly crucial case for deciding between the two theories would come, however, if we could find a form with three consecutive syllables containing the vowel a, the last two of which were unstressed;

the traditional theory would predict that only the last vowel would be affected, becoming u; the preceding vowels would remain unchanged. The rules proposed here, however, would suggest that the three vowels would all be affected, becoming ö-u-u. As was noted above, however, such forms do not appear to exist in the language.

So far then, it appears that all instances of the vowel ö (at least in productive word classes) can be derived from instances of underlying a by the u-umlaut rule. The phonetic value of the first element of the diphthong au (= [öü]) is also accounted for if its underlying representation is /au/. The quality of the second element of this diphthong is produced by rule 2.17, which also accounts for the phonetic value of other instances of u.

II.4 Sketch of the ablaut system. Let us now consider some of the alternations in the stem vowels of strong verbs which are traditionally subsumed under the label ablaut. These verbs are generally divided into seven classes on the basis of the vowel alternations in the reconstructed proto-Germanic parent language, and the division is not always the most felicitous one for the modern languages. Icelandic is one of the few instances of relatively productive use of these alternations, however; much of the original mechanism is still involved rather transparently in the modern language. A complete survey of the strong verb is well beyond the scope of this section (whose purpose is primarily to provide motivation for rule 4.4 below), and the total range of patterns is quite diverse; however, the most common representatives of the original Gmc. verb classes are similar to the following (the principal parts are given in the order

infinitive, preterite singular, preterite plural, past participle):

4.1	a.	bíta	'bite'	beit	bitum	bitinn
		míga	'piss'	meig	migum	miginn
	b.	drjúpa	'drop'	draup	drupum	dropinn
		hnjóta	'stumble'	hnaut	hnutum	hnotinn
	c.	hverfa	'disappear'	hvarf	hurfum	horfinn
		spinna	'spin'	spann	spunnum	spunninn
	d.	bera	'carry'	bar	bárum	borinn
		stela	'steal'	stal	stálum	stolinn
	e.	gefa	'give'	gaf	gáfum	gefinn
		biðja	'ask'	bað	báðum	beðinn
	f.	grafa	'bury'	gróf	grófum	grafinn
		taka	'take'	tók	tókum	tekinn
	g.	leika	'play'	lék	lékum	leikinn
		falla	'fall'	féll	féllum	fallinn

Considering first a verb like hverfa, we see that the structure appears to be [h(v)Vrf] where the V is alternately e, a, u, and o, depending on the tense (and where the y disappears before rounded vowels). Spinna, another verb of class c, shows different vowels in the infinitive and past participle, but these can be accounted for if we note that exactly those verbs whose consonantism is n+C have this shape. Thus, (for strong verbs only, however these are to be marked) we have the rule:

4.2 [- low] → [+ high] / \_\_\_\_\_ n [- cont]

In class c are all strong verbs with the shape C<sub>0</sub>V [+son, +cons] [+cons]. This includes a few with geminate stops (always voiceless), which can be derived by assimilation from an underlying sequence of nasal+voiceless stop: e.g., [drekka] from underlying /drenk+a/. Rule 4.2 does not apply to these verbs, so we assume it applies after the (minor) rule assimilating a nasal

to a following voiceless stop.

The verbs of class b could be assumed to be similar, but with root vocalism jú/jó, au, u, o. Observing that jó appears only before dentals, and jú elsewhere, we can take the latter as basic, and assume underlying jú, au, u, and o. We could further consider the o of the past participle to be derived from u, in order to obtain the more regular series jú, au, u, u. This last step is possible because no strong verbs have u in the past participle except those in which it is derived by rule 4.2 from underlying o. On this basis, we can assume the underlying vowel of the past participles of class c is u as well, and that a rule of inflection applies very early to turn this into o. Rule 4.2 may then apply to turn some of these o's back into u's. Since the vowel u now appears in every form of verbs of class b,<sup>9</sup> we can assume it is part of the root, rather than of the ablaut inflection. Thus, a verb like drjúpa has the basic shape /drVup+a/, where the V position is filled alternately by j, a,  $\emptyset$ ,  $\emptyset$ .

Considering the verbs of class a, we might take them to have shapes such as /bVit+a/ for bfta. The value of V in the third and fourth principal parts is  $\emptyset$ , as was the case with class b; the vowel of the second appears to be e, rather than a, but recall that e can be derived from a by the rule of i-umlaut. Thus, the underlying shape of the diphthong ei is /ai/, and the vocalism of this form is seen to be exactly parallel to the a of classes b and c. What then of the infinitive? If the vocalism were either e or i, we could regularize it with one or the other of classes b and c. Considering the fact that all instances of surface ei can be produced from underlying /ai/, we can ask what would be the result of an underlying /e<sub>1</sub>/ sequence. There are various places in the language where a rule converting e to i before an i in the next syllable is needed. (This rule will not be discussed



here, as its formulation presents considerable difficulty. See Kock (1902) as well as the handbooks for discussion). It can be formulated as an additional environment for the change of rule 4.2; if we assume that it applies to an underlying /ei/ sequence, then /beit+a/ will become /biit+a/, which will be phonetically identical with the required bíta. We have now the basic shapes and vowel alternations of 4.3:

- 4.3    a. /C<sub>o</sub>ViC<sub>o</sub>/, where V = e, a,  $\emptyset$ ,  $\emptyset$   
          b. /C<sub>o</sub>VuC<sub>o</sub>/, where V = j, a,  $\emptyset$ ,  $\emptyset$   
          c. /C<sub>o</sub>VSC<sub>o</sub>/, where V = e, a, u, u

Recalling the rule of u-epenthesis, we see that it could probably be extended to produce the u of the third and fourth principal parts of 4.3c, since /\_\_SC is (approximately) one of the environments of this rule (2.5). If this is so, the epenthesis rule participates in the following ordering relations: it must apply after u-umlaut, since the u which is inserted in forms like 2.2 before the extremely common ending /+r/ does not cause umlaut of a preceding a. An explicit statement to this effect must appear in the grammar, since the ordering epenthesis - umlaut is a feeding order, and the fact that it does not obtain makes the ordering of these two rules marked. Since the rule of u-umlaut is dependent on stress assignment (cf. 2.18), it must follow this latter rule. In forms with u-umlaut, then, the three rules will have to apply in the order stress - umlaut - epenthesis. In the roots of the strong verbs, however, the vowel inserted by the epenthesis rule must receive stress, and hence the epenthesis rule must precede the stress rule. Since u-umlaut does not apply in these forms, however, there is no problem (on the assumptions made here); the only marked ordering involves umlaut, and the remaining orderings follow from general principles. Thus, the fact that the rules of stress assignment

and u-epenthesis apply in different relative orders depending on whether u-umlaut applies in the form follows from the one ordering statement "umlaut precedes epenthesis". Such a situation would be impossible in a grammar operating under a constraint of linear ordering, of course, and the generalization of the two epenthesis processes would have to be given up.

At this point, only the j of the infinitive of class b prevents us from considering all of the first three verb classes as instances of the same alternation e, a, Ø, Ø, with the surface deviations from this predicted by the phonological shapes of the roots and a set of reasonable rules. In fact, the vowel e never appears as the first element of a morpheme-internal diphthong unless it is the result of applying i-umlaut to an underlying a. If the underlying vowel of the infinitive of class b were taken to be e, we could assume a rule which converted this e to j, with no ill effects on the rest of the grammar. It would be necessary to shift the stress from the e (where it would be assigned by the basic stress rule) to the following vowel; in fact, this rule will be sufficient by itself, for the resultant unstressed e will be reduced to i by the vowel reduction rule, and then a later rule (to be discussed below) will convert this to j before a stressed vowel. We are now in a position to reduce all of classes a, b, and c (the most common classes) to alternations of e, a, Ø, Ø. Another rule to account for the unexpected á of the preterite singular of class d (whose basic shape is C<sub>0</sub>VS) allows us to extend this schema to this class, as well. Class e is more complex, but at least the vocalism e, a of the first parts is regular. All of this lends some support to rule 4.4 below, since it is this that allows us to regularize most of the principal parts of the first five classes of strong verbs. These are by far the largest and most productive classes.

4.4<sup>10</sup>

$$\begin{array}{ccc}
 \left[ \begin{array}{c} e \\ + \text{ stress} \\ \alpha \text{ tense} \end{array} \right] & [+ \text{ syll}] \longrightarrow & \left[ \begin{array}{c} 1 \\ - \text{ stress} \\ - \text{ tense} \end{array} \right] \left[ \begin{array}{c} 2 \\ + \text{ stress} \\ \alpha \text{ tense} \end{array} \right] \\
 1 & & 2
 \end{array}$$

Some additional support for rule 4.4 can be derived from the declension of nouns with final é, such as fé, 'money'. The genitive singular of this noun (formed by the addition of the suffix /-ar/) is fjár, exactly as would be expected if rule 4.4 applied to the representation /fé+ar/. Other forms with final tense é, such as tré 'tree', hné 'knee', hlé 'shelter', etc., have dative plurals with the vocalism já: trjám, hnjám, hljám, etc. If we assume these roots end in lax e, this vowel will automatically be tensed in final position, as in the nom. sg.; Icelandic words cannot end in stressed lax vowels. Thus, the dative plural will have, e.g., /tre+um/ as its underlying representation. The e will be broken to ea, as detailed in the next section below, stress will shift to the a, and the resultant sequence a+u is identical with á (if we assume, as we must, that u-umlaut is blocked if the vowels are adjacent over a +boundary). Further details of the role of rule 4.4 in the process of breaking will be seen below.

II.5 Breaking. The process of breaking, which has just been alluded to, is one of the most popular topics in historical Scandinavian phonology. It consists in an alternation between lax, stressed e and the sequence ja or jö. The sequence ja is found, roughly, before lax a in the following syllable; jö is found before u in the next syllable. It is not possible to give a complete formulation of the rule, but the presence of a lax back

vowel in the following syllable is the main conditioning factor:

- 5.1 a. gjöf (from /gef+u/; cf. gefa 'to give') 'gift'  
       gen. sg. gjafar (from /gef+ar/)  
       dat. pl. gjöfum (from /gef+um/)
- b. gjalda (from /geld+a/) 'to pay'  
       geld 'I pay'  
       gjöldum 'we pay'

It is clear that the difference between ja and jö is another reflex of the operation of the u-umlaut rule, and that in fact the breaking rule itself is only responsible for the change e to ja. Keeping the rule 4.4 above in mind, it is possible to formulate the change simply as the epenthesis of an a after the e; after this, rule 4.4 will shift the stress from the e to the a, and the e will be reduced to j as was the case with the e of, e.g., drjúpa in the strong verb system. This provides further evidence for the correctness of rule 4.4, since it allows the much more natural statement of the rule of breaking:

$$5.2 \quad \emptyset \rightarrow a / \left[ \begin{array}{c} e \\ + \text{stress} \end{array} \right] \text{ --- } C_o \left[ \begin{array}{c} + \text{syll} \\ + \text{back} \\ - \text{tense} \\ - \text{stress} \end{array} \right]$$

(with some further conditions)

There has in fact been a controversy in the scholarly literature on the subject over the correct formulation of the process of breaking. In particular, it has been suggested that the vowel which this epenthesis process should insert should not be a, but rather another vowel with a different quality, in particular, an unrounded mid back vowel [ə]. While the formulation just given, where a is inserted, is much more appealing at first sight, since the vowel inserted behaves in every way as an a in Mod.

Icelandic, there are facts that argue against this formulation of the change at the point at which it originally applied in Old Norse. Since it is only the O.N. change that is ever discussed (in light of the general neo-grammarian *cum* taxonomist bias of this literature), the issue of the form of the Mod. Icel. rule is rarely raised.

Benediktsson (1963a) has given the relevant facts, which are as follows: the basic vowel ø in general, when it appears in the small class of environments in which lengthening of back vowels took place (i.e., before l+non-dental C, before h+C), gives simply long ô, whose modern reflex is long á. In the case of the products of breaking, however, where we would expect to find jô, which would give já, we find instead jó. If we assume the e became ea, then ja, and then jô, and then was lengthened, as the treatment above would assume, there is no way to predict, e.g., the vowel in mjólk (from /melk+u/) by breaking, umlaut and lengthening. On Benediktsson's analysis, however, where the inserted vowel is [ə], this follows naturally: e becomes eə, which becomes jə; the umlaut rule then rounds this [ə] to [o] (by a natural extension), and lengthening can now apply to this vowel, giving the correct output ô. If lengthening does not apply, and the vowel [ə] receives the stress (as of course it will, by 4.4), the resultant vowel is identified with a by a rule such as:

5.3

$$[+ \text{ back}] \longrightarrow [+ \text{ low}] \quad / \quad \left[ \begin{array}{c} \text{---} \\ - \text{ round} \\ + \text{ stress} \end{array} \right]$$

Such a rule is perfectly natural; a relation between a and [ə] is found in many other languages. The only problem with this solution is the fate of unlengthened but umlauted [ə]: there is no obvious reason why this should be identified with ô, as was in fact the case, rather than with o. Benediktsson points out that there were no other sequences of j+back rounded vowel

in the language (at least with lax vowels), and that there was thus no reason why phonetic [o] should not be identified with the taxonomic phoneme /ö/ in this position. True, there is no reason why not, but then again, there is no reason why the identification should be made; the phonetically natural identification would be with /o/, apparently, and anything else seems rather like the identification of English [h] and [ɣ] as one taxonomic phoneme.

When we recall the rule of vowel shift proposed above (rule 2.13 above), however, we see that there is a relation already established between ö and o, if such a rule is in the grammar. Assuming that this rule was part of Old Norse (perhaps together with a version of the fronting rule, 2.17, applying only to mid vowels, which we know to have been part of the grammar at some time), we could obtain the correct results by adopting Benediktsson's solution, together with some mechanism for preventing the vowel shift from applying to the output of breaking. Such a device might be, for instance, the requirement "vowel shift precedes stress shift" (= rule 4.4), together with the restriction of vowel shift to stressed position. Other devices can be envisioned whose effect would be the same; assuming this one, we would have the following derivations (where stress is marked with ´, and length with ¯):

5.4	/..élku../	/..épu../	/..épa../	/..ápu../
(breaking)	..é <sup>o</sup> lku..	..é <sup>o</sup> pu..	..é <sup>o</sup> pa..	
(umlaut)	..éolku..	..éopu..		..ópu..
(V-shift)				..ópu..
(stress-shift)	..jólku..	..jópu..	..j <sup>o</sup> pa..	
(rule 5.3)			..jápa..	
(lengthening)	..jólku..			
(fronting)		..jó <sup>o</sup> pu..		..ó <sup>o</sup> pu..
output:	[..jó <sup>o</sup> lku..]	[..jó <sup>o</sup> pu..]	[..jápa..]	[..ó <sup>o</sup> pu..]

With this device, we see that the phonetic result of not lowering the rounded [ə] is to make it identical with the (shifted) ö from underlying /a/. This then gives us further evidence for the claim that the vowel shift rule (2.13) has in fact been part of the grammar for some time. This assumption allows us to give a reasonably natural formulation of the results of the lengthening rule when applied to the product of breaking.

The above discussion assumes that we accept a version of Benediktsson's formulation of breaking as a description of the original change. Notice, however, that this is crucially related to the fact that it is necessary to explain an apparent alternation between long and short products of the rule. In fact, the lengthening rule which caused us to posit such a solution was always highly limited in its application; it affected primarily a handful of relatively isolated forms, whose new (lengthened) values were presumably entered in the lexicons of new speakers of the language with their new values, thus obviating the lengthening rule itself as a productive part of the phonology. Once this was done, new speakers were no longer driven to construct the grammar underlying the derivations in 5.4; in the absence of length alternations which would require this, it would be possible to dispense entirely with the rule 5.3 by reformulating the breaking rule to insert a rather than [ə], as is suggested in the formulation of this rule as 5.2. Thus, the lexicalization of the effects of the lengthening rule led to a restructuring of the grammar from that proposed by Benediktsson to the more natural form assumed here, which is the form of the rules in Modern Icelandic reflecting the change. This is an excellent example of the differences which can arise between the rules of a synchronic grammar which describe a given alternation and the rules which describe the change which originally gave rise to the alternation.

II.6 The i-umlaut rule. Recall the forms in 1.3 above, which were presented as examples of the alternation known as i-umlaut. Besides the comparative forms of many adjectives and certain derivatives, this alternation can be observed in the present versus the infinitive of strong verbs, the present versus the preterite of class i weak verbs, the datives of a few strong nouns (whose ending represents underlying /-f/, rather than /-e/) and at various other places in the language. The surface correspondances are as follows:

6.1	basic vowel	umlauted vowel
	a	e
	á	æ
	o	e
	ó	œ
	u	y (=i)
	ú	ý (=f)

Note that the change of e to i, mentioned above in connection with the ablaut system, is not a part of this process, even though it is also conditioned by a following i. This is shown by the fact that it is much more restricted; e.g., in highly productive umlauting positions such as the present indicative of strong verbs, it does not generally apply.

Consider the form this rule must take. It is, of course, primarily an alternation between front and back vowels. This front/back alternation pattern seems to be destroyed, however, by the vowel æ, corresponding to both á and ó. This is phonetically [aj], with a back vowel nucleus. There is reason to believe, however, that this vowel is a front vowel at at least some point in its derivation: there is a rule in the language, already alluded to above, whereby velars are fronted (i.e., palatalized) in the position before a following front vowel. This rule applies before æ as well as before phonetically front vowels; this would lead us to believe that the umlaut



process should produce a front vowel [ $\bar{æ}$ ] which is later backed after the application of the velar fronting rule. Additional support for this assumption will be provided in the next section; there it will be shown that the glide associated with this vowel will follow as an automatic consequence of the assumption that it is [ $\bar{æ}$ ] at some point before other rules apply.

Thus, the umlaut rule can be made to make all of the vowels [- back]; should it also make them [- round]? If we simply make everything [- back, - round], we will still have to raise the result of umlauting a ([ $\bar{æ}$ ]) to e, which will be no problem, due to the absence of any other such vowels in the language, but it will also be necessary to lower to [ $\bar{æ}$ ] the result of umlauting ó, with the result that all tense é's would be lowered. In fact, there are other vowels in the language (é) which should not be affected. Evidently then, if the unrounding is to be accomplished as part of the rule of i-umlaut, the height changes in the non-high vowels produced must also be part of the rule, in order to prevent the opposition between é and æ from being neutralized (as it was in Faroese; cf. chapter III). We could thus propose the following formulation of the structural change of the rule:

6.2

$$[+ \text{back}] \rightarrow \left[ \begin{array}{l} - \text{back} \\ - \text{round} \\ \langle \alpha \text{ low} \rangle \end{array} \right] / \left[ \begin{array}{l} \text{---} \\ + \text{syll} \\ \langle - \text{high} \rangle \\ \alpha \text{ tense} \end{array} \right] / \text{--- E}$$

This is hardly the world's most natural rule; however, it is apparently the correct description of the facts. If so, it would be a good example of the way in which a sequence of perfectly natural rules (for which see Benediktsson, 1958) can result in a restructured grammar containing an unnatural rule. Each of the individual rules that are responsible for the evolution of the

modern Icelandic vowel system is, in itself, perfectly natural; the result, however, is the rule 6.2.

It has already been noted that one environment for the operation of rule 6.2 is the position before an i in the following syllable. The rule also applies to the vowel a, making it e, in the position before a velar followed by e. For instance, the dative singular of dagur is degi, from underlying /dag+e/. The fact that it is precisely a velar followed by a front vowel which serves as the environment for the rule is suggestive; it will be recalled that in this position, a velar is fronted, acquiring the features [+high, -back]. These are, of course, precisely the features that distinguish the vowels that cause umlaut (i.e., i and e) from those that do not. This suggests that, although the process affects only one vowel, it has considerable similarity to the rule that applies before vowels. We are thus tempted to formulate the environment for 6.2 as follows:

$$6.3 \quad (\text{rule 6.2}) / \left[ \begin{array}{c} \text{---} \\ \langle \langle \begin{array}{c} - \text{ low} \\ + \text{ tense} \end{array} \rangle \rangle_a \end{array} \right] \langle \langle \begin{array}{c} c \\ o \\ b \end{array} \rangle \rangle_b \left[ \begin{array}{c} + \text{ high} \\ - \text{ back} \\ \langle \begin{array}{c} c \\ cons \end{array} \rangle_c \end{array} \right]$$

Condition: if a, then b and c

This rule applies before i or e in the following syllable, or before an immediately following fronted velar if the vowel affected is a.

A problem arises with this formulation; the vowels that are produced by umlaut (6.2 cum 6.3) are front vowels; as such they cause velar fronting. Thus, velar fronting must follow umlaut; but if umlaut is at least partially conditioned by the frontness of a following velar, the velar fronting rule must also precede umlaut. Both of these are feeding orders, however, and the result is that the correct results will be obtained (on the assumptions adopted here) simply by not stating an ordering constraint.

II.7 The resulting vowel system. If we consider the vowels whose orthographic representations include an acute accent (á, ó, ú, é, í), we see that with the exception of é, each consists of the corresponding simple vowel followed by a high glide of the same backness. Thus, the rule which converts these vowels (whose representation we have suggested above is the same as that of the lax vowel, but with the diacritic feature [+ tense]) into their phonetic outputs should create just such a glide. Let us assume it does so by producing a [- low] copy of the vowel following it, and that these vowels, being unstressed and adjacent to another vowel, are then converted to glides by another rule. This other rule will be justified below; the vowel copying rule is approximately as follows:

$$7.1 \quad \emptyset \longrightarrow \left[ \begin{array}{l} + \text{ syll} \\ - \text{ low} \\ \alpha \text{ high} \\ \beta \text{ back} \\ \beta \text{ round} \end{array} \right] / \left[ \begin{array}{l} + \text{ syll} \\ + \text{ tense} \\ \alpha \text{ high} \\ \beta \text{ back} \\ + \text{ stress} \end{array} \right] \text{ — }$$

From underlying á, ó, ú, í, this will give the following:

$$7.2 \quad \begin{array}{lll} \acute{a} & \longrightarrow & \acute{a}o \\ \acute{o} & \longrightarrow & \acute{o}o \\ \acute{u} & \longrightarrow & \acute{u}u \\ \acute{i} & \longrightarrow & \acute{i}i \end{array}$$

Such a process is of course perfectly natural from the historical point of view; in fact, the tense vowels developed from original long vowels, which were double length versions of the short vowels. The product áo of underlying long á is supported by the present day reflex of this vowel in the Eastern Scandinavian dialects: ä ([ɔ] or [o]).

Observe that, if the i-umlaut of á and ó, æ, is taken to be [æ̥], a front vowel, 7.1 will apply to it to give [æ̥e]. The e will later become j, and thus the glide on this vowel follows as a matter of course from the

assumption it is a front vowel.

Consider now the effects of applying 7.1 to the underlying tense vowel é: the result will be ée. But now notice that the stress shift rule (4.4) will apply to this sequence, converting ée to eé. The glide rule will then convert the first member of the sequence (the unstressed one) to a glide, rather than the second element of the sequence, as with other underlying tense vowels. This will give (correctly) the result [je] for the underlying vowel é, a value which must otherwise be regarded as highly unnatural. The ability of this analysis to explain such an apparent irrationality of the vowel system is evidence of the strongest possible sort for the correctness of the stress shift rule (4.4) and the diphthongization rule (7.1).

It was noted earlier that i's which do not cause umlaut can be represented as underlying e, and that the feature of [ $\pm$  tense] is available to distinguish, among those i's that do cause umlaut, between those that appear phonetically and those that are deleted in certain positions. Parallel to forms like hnot/hnetur 'nut/nuts', which require an i in the plural which disappears from the phonetic representation, there are forms like those in 7.3 below, which require a (final) u that is deleted from the phonetic representation:

- 7.3    a. meðal 'medicine'; pl. meðal, ul  
          b. barn 'child'; pl. börn  
          c. gjöf 'gift' (from /gef+u/; cf. gen. sg. gjafar)

Again, the same feature of [ $\pm$  tense] is available to differentiate the disappearing from the non-disappearing u's.

The majority of the deleted lax high vowels are found in final position. Forms such as bök/bækur (= /bök+i+r/) 'book/books' show that a final r does not prevent the vowel from being deleted; a form like the genitive söngs (from /sang+u+s/) 'of the song' shows that a final s does not prevent the deletion,

either. Before some vowels, however, these same vowels are not deleted: the nom. pl. of söngur (from /sang+u+r/), whose genitive sing. is söngs, is söngvar, with the theme vowel u appearing as y before the ending /-ar/. The relation between w and y is sufficiently widespread for us to assume that this y represents an intermediate stage of w from the underlying u, exactly parallel to the j from underlying i.

Some final consonants do prevent the deletion of lax unstressed high vowels, however. In the dative plural ending -um, the vowel must be assumed to be a lax u, since it causes breaking in a preceding e. Let us assume, then, that only dentals may intervene between the vowel to be deleted and the word boundary:

$$7.4 \quad \left[ \begin{array}{l} + \text{ syll} \\ + \text{ high} \\ - \text{ tense} \\ - \text{ stress} \end{array} \right] \rightarrow \emptyset / \text{ \_\_\_\_\_\_ } [+ \text{ coronal}]_0 \#$$

After this rule applies, the rule of vowel reduction (2.14) applies. All remaining unstressed vowels which stand in position before another vowel can be made glides: any remaining vowels which follow other vowels (which is only possible if the first vowel was preserved as a vowel by being stressed) then become glides, such as the second elements of the diphthongs in 7.2. These two rules are as follows:

$$7.5 \quad \begin{array}{ll} \text{a. } [- \text{ stress}] \rightarrow [- \text{ syll}] / \text{ \_\_\_\_\_\_ } [+ \text{ syll}] \\ \text{b. } [- \text{ stress}] \rightarrow [- \text{ syll}] / [+ \text{ syll}] \text{ \_\_\_\_\_\_ } \end{array}$$

These two rules are obviously related; the mirror image notation, which will be discussed in some detail in chapter IV below, allows us to collapse them into the single schema 7.6:

$$7.6 \quad [- \text{ stress}] \rightarrow [- \text{ syll}] \% \text{ \_\_\_\_\_\_ } [+ \text{ syll}]$$

Not all of these glides appear in the output, however; they are retained only after stressed syllables which either end in a velar or are composed of a lax vowel followed by exactly one consonant. We can thus write the following rule to delete the other glides:

$$7.7 \quad \begin{bmatrix} - & \text{syll} \\ - & \text{cons} \end{bmatrix} \rightarrow \emptyset / \left\{ \begin{array}{l} \begin{bmatrix} + & \text{syll} \\ - & \text{stress} \end{bmatrix} \quad c_0 \quad \text{---} \\ \begin{bmatrix} + & \text{syll} \end{bmatrix} \quad c_2 \quad \text{---} \\ \begin{bmatrix} + & \text{cons} \\ - & \text{high} \end{bmatrix} \quad \text{---} \end{array} \right\}$$

Examples of the presence and absence of these glides, which correspond to the underlying theme vowels, are:

- 7.8 a. stöð 'station'; pl. stöðvar (stem /stað+u/; lax vowel followed by one consonant)
- b. pöntun 'order'; gen. sg. pöntunar (stem /pantan+u/; unstressed vowel)
- c. veggur 'wall'; gen. sg. veggjar (stem /vagg+i/; stem ending in a velar)
- d. lækni 'doctor' pl. lækna (stem /lákni/; stem ends in more than one consonant)

These rules appear to permit us to assume the presence of a high vowel (front for i-umlaut, back for u-umlaut) in all of the forms in which one of the umlaut operations applies. The great majority of glides can also be derived from underlying vowels; perhaps, in fact, the glides can be entirely eliminated from the underlying segment inventory.

The rules which have been proposed in this chapter can only be assumed to be valid for forms that belong to classes which take part in phonological alternations. Nouns, verbs, and adjectives all belong to such productive classes; it seems reasonable to argue that in these words, even if the

particular form does not alternate, it should be assigned an abstract representation and derived from this by the productive rules of the grammar. It was argued above in section 3 that this was necessary, at least for instances of ö which could (but did not have to) be derived from underlying a by u-umlaut. In invariant forms, however, such as the adverb mjög 'very' (which could be derived from /meg+u/), it may well be that the surface form is very nearly the correct lexical form, rather than the more abstract form which would lead to the same phonetic output. Similar remarks apply to the vowels in any form which are in such a position as to be impervious to the effects of any possible morphological process (and hence could not participate in any alternations). Kiparsky's observations (cf. Kiparsky, 1968a, 1968b) concerning the development of such unproductive forms in German dialects form the basis of this remark; it is still necessary to find empirical evidence within Icelandic for such a process of lexicalization.

In morphemes belonging to productive classes, however, we have arrived at the following vowel system (including the glides):

7.9	[+ tense]		[- tense]		diphthongs	
[+ high]	í		ú	i	u	ai au
[- high]						
[- low]	é		ó	e	o	
[+ low]						
	- +	á	- +	- +	a	- +
	[ ± back] [ ± round]		[ ± back] [ ± round]			

## Notes

<sup>1</sup> Stress in Icelandic is basically on the first syllable of a word, and on the first element of a compound. Secondary stress appears on other vowels sometimes, however, especially in loan words. Many Icelandic morphemes are monosyllabic. In those few forms showing the effects of secondary and lower stress (all of which, together with primary stress, count as [+ stress]), I have simply assumed its placement.

<sup>2</sup> Cf. Benediktsson, 1963b, 1969.

<sup>3</sup> This glide follows the vowel in á, ó, ú, and í, but precedes it in é. The basis for this fact will be discussed below in section II.7. The one other way in which the statement in the text is not strictly accurate is in connection with the relationship between ú and u. The lax member of this pair is phonetically [ü], while the nucleus of the tense member is [u]. This difference is due to a fairly late rule which will be brought up again below.

<sup>4</sup> In a great many forms, tenseness could be assigned by a morpheme structure rule not unlike rule 1.2, though this is not true of many other forms. It is unclear to what extent generalizations can be extracted from the distribution of tenseness in the modern language. Something like rule 1.2 has always been part of the grammar; such a rule originally assigned vowel length, a feature which was later interpreted as distinctive tenseness after the diphthongization of the original long vowels. At this point, a new rule (1.2) arose, re-introducing the non-distinctive length. A more accurate description of these facts will be found in Benediktsson (1968).



<sup>5</sup> For certain classes of borrowed words, I assume a lexical representation involving a stressed vowel. When the normal stress rule applies to such forms, assigning primary stress to the first syllable, this will lower the stress on the lexically marked syllable to secondary. A late phonetic rule eliminates all but the highest stress within a span delimited by # boundaries. While less than perfectly esthetic, this solution is still quite plausible. In the first place, the syllables which must be marked in the lexicon as stressed are generally ones which have some stress in the language from which they are borrowed; for example, Almanak, in which we assume a lexical stress on the final syllable, is an obvious German borrowing; in German it also has at least some stress on the final syllable, as shown by the failure of this syllable to reduce. It is in fact a fairly common device in the Germanic languages, where there is a more or less constant trend toward initial stress, to mark a word as foreign by stressing the final syllable. In almanak, the stress on the final syllable is responsible for the fact that in the dative plural, almanökum, not only does the vowel not reduce, but the umlaut does not affect the first and second vowels. Another instance in which lexical stress plays a role is in the puzzling diversity of possible dative plurals for the word banani. The expected form, bönunum does occur, but some speakers also have the form banönum. This is what would be expected if the form had a stress on the second syllable, as of course it does in the source language. Some speakers, in addition, have a third form of the word: banunum. This is a totally isolated shape: it could be derived, however, by assuming again that the form has secondary stress on the second syllable, as required to get banönum, but that the vowel undergoes reduction anyway (irregularly) by exceptionally allowing the reduction rule to apply after the stress erasure rule. Such an ordering would be a feeding order, a relaxation of a

normally enforced non-feeding order. Lexical exceptions of exactly this type would be expected on the assumptions adopted here; for some further discussion, cf. chapter VI. The particular exception under discussion (the reduction of the second syllable of banunum) could be motivated either by the increasing assimilation of the word (as indeed Iceland has increasingly assimilated the banana, which they grow quite successfully in hothouses heated by spring water), or by the strong feeling among native speakers that the vowel ö does not occur in foreign words.

<sup>6</sup> It should be pointed out that ölnum is only one possible form of this word; more common is the (irregular) álnum.

<sup>7</sup> This stanza can be glossed (approximately) as follows:

"It can be seen that we were  
rejoicing in battle with Olaf;  
I got from it, Hilda of the horn,  
a wound, and little peace.  
(The light) shines from my shield  
the storm grows too cold for the skald;  
the archers have nearly made me left-handed."

The sensitive should not cringe at this translation; Skaldic poetry is largely untranslatable.

<sup>8</sup> Cf. Halle, 1967; Halle and Keyser, 1967.

<sup>9</sup> Some of these u's (and o's) are tensed; this is a consequence of a largely ad hoc rule which, I assume, assigned the tenseness first to the preceding segment, from which it is later transferred to the u. This is one of

the few examples in the language, incidentally, of an alternation between a tensed and a corresponding lax vowel.

<sup>10</sup> Though I have given no evidence to support the shift of tenseness, some examples will appear below.

### Chapter III - The Vowel System of Faroese

III.1 Introduction. Faroese, the West Scandinavian language spoken in the Faroe Islands, is closely related to Icelandic in remaining quite close in many ways to Old Norse. One way in which it differs both from Icelandic and from Old Norse, however, is in its surface vowel system, whose description has produced no little discussion among Scandinavianists. Most of the facts concerning the interrelationships among the various elements of the phonological system are well described in the traditional grammars (Hammershaimb, 1891; Krenn, 1940; Lockwood, 1955), historical accounts (Haegstadt, 1917; Rischel, 1968), and dialect surveys (Werner, 1964 and 1968) as rules for the pronunciation of the standard orthography, or as a sequence of historical events (sound changes) which have resulted in the modern language. These two aspects of description are closely interrelated, since the orthography is a consciously archaic one, which was designed to preserve as much of a resemblance to the other West Scandinavian dialects and to Old Norse as possible. Despite the fact that other orthographies have been proposed, including ones based on the surface forms of words to a much greater extent than the standard, the populace has shown little interest in change. The aptness of the apparent archaism will perhaps become clearer below.

Most modern treatments of Faroese take the position that the orthography is too artificial to be taken seriously, and have rejected the notion that it might provide a basis for any insightful observations. Treatments such as Bjerrum's (1949; a glossematic approach), and Roe's (1965; a straightforward taxonomic phonemic description) have simply started from the phonetic facts and extracted some minimal regularities from them without attempting to account for the role that the various elements posited play in phonological

alternations. O'Neil (1964; and McCawley, 1968, which is simply a reformulation of part of O'Neil's description) has adopted a more abstract approach, admitting some facts concerning the interrelation of the elements of the vowel system and a non bi-unique relation between the phonemic and the phonetic levels. This treatment too, however, seems to have erred in staying too close to the phonetic facts, thereby missing some interesting generalizations.

III.2 Inventory of vowels. Let us first consider the vowels which are represented in the orthography as unaccented, simple vowels. Forms containing these vowels show a consistent alternation between two sets of phonetic values, depending on whether the syllable in which they appear is long or short:<sup>1</sup>

2.1	<u>a</u> :	lamin	[l <del>a</del> amIn]	'lame, nom. masc. sing.'
		lamnir	[lamnIr]	'id., nom. masc. pl.'
	<u>e</u> :	gera	[ʎe:ra]	'to do'
		gert	[ʎert]	'(thou) dost'
	ø :	kjøti	[čö:ti]	'meat, dat. sg.'
		kjòts	[čöts]	'id., gen. sg.'
	<u>o</u> :	tora	[to:ra]	'to dare'
		tordi	[tɔrdi]	'dared'
	<u>i/y</u> :	ymissur	[i:mIssUr]	'different, nom. masc. sing.'
		ymsir	[ImSir]	'id., nom. masc. pl.'
	<u>u</u> :	rugur	[ru:wUr]	'rye, nom. sg.'
		rugbreyð	[rUbbrej]	'ryebread'

Assuming that these vowels have the feature values given in table 2.2 below, we can then account for the alternation shown in 2.1 with a rule which specifies vowels as lax in short syllables, and tense elsewhere.

The phonetic reflexes of tenseness will then include length, in the case of the non-low vowels; in the case of low a, tenseness will be realized as diphthongization to [æ̯a]. In at least some dialects, tense e and o are raised to i and u before a, by still another rule. A further dialectal phenomenon is the change of lax a to [ɛ] before a velar nasal.

## 2.2

vowel	[+back]	[+high]	[+low]	[+round]
a	+	-	+	-
e	-	-	-	-
ø	-	-	-	+
o	+	-	-	+
i/y	-	+	-	-
u	+	+	-	+

Turning now to the vowels represented in the orthography with an acute accent, we see that they are phonetically diphthongs, with the following values in long syllables:

2.3	f/ǝ	: [uj]
	ú	: [uɥ]
	ó	: [ɛw], [öw], or [ɔw], depending on dialect
	á	: [oä]

Recalling the phonetic value of tense a ([æ̯a]), we see that it has been diphthongized from its underlying value of [a] by the insertion of a segment [æ̯] before the vowel. Let us assume that the diphthongs in table 2.3 are also created by a rule of "regressive diphthongization" from underlying representations as the corresponding simple vowels with some diacritic feature such as [<sup>±</sup> long]<sup>2</sup> to differentiate them from the short vowels. Assuming that such a rule of regressive diphthongization applies to the marked /f, ú, ó, á/, it might look like 2.4:

$$2.4^3 \quad \emptyset \rightarrow \left[ \begin{array}{l} + \text{syll} \\ - \text{low} \\ \alpha \text{ high} \\ -\beta \text{ back} \\ -\beta \text{ round} \end{array} \right] / \text{---} \left[ \begin{array}{l} + \text{syll} \\ + \text{long} \\ \alpha \text{ high} \\ \beta \text{ round} \end{array} \right]$$

This will create the diphthongs /ui, iu, ɛo, oa/. Let us now assume that a rounding adjustment rule applies, which rounds a high vowel before a following rounded vowel; such a rule would be restricted to the position before a high rounded vowel in those dialects that have the value [ɛw] for ó, rather than [øw] or [ow]. All diphthongs in Faroese are falling (as will be seen later, the segments in question do not represent the only diphthongs in the language), and it is reasonable to assume a process that converts the second element of a vowel sequence into a glide. The effect of these processes, then will be to convert underlying [+ long] /ú, í, ó, á/ into [üw, uj, ɛw or øw, oɔ].

Up to this point, we have been assuming that only the vowels marked with accents in the orthography are affected by 2.4: that is, that the long vowel system is like this:

$$\begin{array}{rcl} 2.5 & /i/ & /ú/ \\ & & /ó/ \\ & & /á/ \end{array}$$

This system lacks a tense mid front vowel /é/. Historically, this came about as follows: the vowels of 2.5 are etymologically long vowels; long é was merged with low front æ (the umlaut product of long á), which was then shortened to æ, which merged with the back vowel a. The result was the loss of long é as a distinct vowel. Observe what would be the effect of positing an underlying /é/, however; rule 2.4 would convert this to /oe/, and this latter would then become [oj]. Such a phonetic diphthong does in fact exist in Faroese, corresponding to orthographic oy. This diphthong, then,

could be eliminated from the underlying representation and replaced with the more natural segment /é/, which would complete the long vowel system.

Evidence for or against this solution is extremely difficult to find. There are virtually no alternations in Faroese between underlying long vowels and the corresponding short vowels; it is, therefore, not at all an argument against this proposal that oy does not alternate with e in any form. On the other hand, there is no structural reason to represent oy as underlying /oj/. There are two other diphthongs in the language which have not yet been discussed: ei (= [aj]) and ey (= [ej]), whose role in alternations such as the ablaut system provides compelling evidence for their representation as underlying /ai/ and /au/.<sup>4</sup> The diphthong oy, however, does not play any comparable role in alternations, and is thus an isolated element of the system.

Having discovered that there is no reason not to represent oy as underlying /é/, we can ask further whether it is plausible for such a representation to have arisen, and whether there are any historical facts which it assists in explaining. This historical question is of course completely independent of the synchronic representation, but answers to it can at least be suggestive. Historically, Faroese oy represents an Old Norse diphthong variously represented as öy, ey, etc., which originally arose as the i-umlaut of the diphthong au. In Icelandic, the reflex of this diphthong is [ej], with which the diphthong ei (= /ai/) has fallen together. In the Faroese of ca. 1500, the diphthong oy seems to have been phonetically [ej], with ei (= mod. [aj]) represented as [æj], if one is to credit the scanty documentary evidence for the period. It would not, then, have been implausible for the isolated diphthong [ej] to have been treated as a simple long vowel /é/ at the time the diphthongization rule 2.4 was added to the language. The other long



vowels were apparently similar in nature at the time;  $\underline{i}$  = [ij],  $\underline{u}$  = [uw],  $\underline{o}$  = [ow], etc. At this stage of the language, then, it would be natural for children learning the language to have constructed a grammar in which [ej] was represented as /é/, a vowel which had otherwise disappeared from the language as described above. When they then added the diphthongization rule 2.4 to their grammars, this /é/ would be treated in the same way as the other original long vowels, and converted to [oj].<sup>5</sup> Thus, the assumption that this diphthong was re-analyzed as underlying /é/ at a very early point provides an explanation of its otherwise inexplicable development. As such, we have supporting evidence for both the representation of oy as /é/, and the rule of diphthongization, 2.4.

One further rule should be mentioned at this point to explain the phonetic shapes of vowels: diphthongs ending in [ $\underline{a}$ ] (i.e., [ $\underline{æ}\underline{a}$ ] from /a/ in long syllables, and [ $\underline{o}\underline{a}$ ] from /á/) lose this element with compensatory lengthening before an immediately following [a]:

$$2.6 \quad \begin{array}{ccccc} [+ \text{syll}] & \underline{a} & a & & \\ & 1 & 2 & 3 & \longrightarrow \quad \begin{array}{ccc} 1 & \emptyset & 3 \\ [+ \text{tense}] & & \end{array} \end{array}$$

We have now arrived at the following vowel system for the underlying representations of Faroese forms:

2.7	<u>short vowels</u>			<u>long vowels</u>		<u>diphthongs</u>	
	i		u	í	ú	ai	au
	e	ø	o	é	ó		
		a			á		

It will be observed that, with the exception of  $\underline{o}$ , this is the same system as that arrived at for Icelandic.

III.3 The glide-epenthesis rule. There is a rule in Faroese that inserts a glide between two adjacent vowels in forms like 3.1:

- 3.1
- a. siður [si:jUr] 'cuscum'
  - b. kvæði [kvæ:ji] 'ballad'
  - c. deyður [dej:jUr] 'dead'
  - d. soðin [so:jIn] 'boiled'
  - e. maður [mæ:vUr] 'man'
  - f. leður [le:vUr] 'leather'
  - g. suður [su:wUr] 'south'
  - h. húðir [hü:wIr] 'skins'

We note that the glide inserted is one of j, w, or y, depending on the surrounding vowels. All instances of y are following a [- high] vowel; furthermore, no instances of w appear in such position. It is a general property of the language that w becomes y in all positions except after a [+ high] vowel. We can assume this rule here, and will see other instances of its application below. Thus, the glide insertion process which we presume to be responsible for the forms in 3.1 need only discriminate between environments in which j is inserted and those in which w is inserted.

Table 3.2 gives the glides which are inserted to break up any given sequence of vowels  $V_1V_2$ . Since  $V_1$  is the only one of the pair that could possibly be stressed, and only the three vowels a, i, and u can appear in unstressed positions, these are the only values that need be considered for  $V_2$ :

3.2

$V_1 \backslash V_2$	a	i	u
i	j	j	j
u	w	w	w
e	∅	j	w
ø	∅	j	w
o	∅	j	w
a	∅	j	w

We see that the glide insertion process can be described by a pair of rules, one of which inserts a 'similar' glide after a [+ high] vowel and before another vowel, and another which inserts a 'similar' glide after any vowel and before a [+high] vowel. Thus, if neither of the two vowels  $V_1$  and  $V_2$  is [+ high], no glide at all gets inserted; and the quality of the first vowel, if it is [+ high], takes precedence over that of the second. These two rules can be formulated as:

$$\begin{array}{lcl}
 3.3 \text{ a.} & \emptyset \rightarrow & \begin{bmatrix} + \text{ son} \\ - \text{ cons} \\ - \text{ syll} \\ + \text{ high} \\ \alpha \text{ back} \end{bmatrix} / \begin{bmatrix} - \text{ cons} \\ + \text{ high} \\ \alpha \text{ back} \end{bmatrix} \text{ \_\_\_\_\_\_ } [- \text{ cons}] \\
 \\
 \text{b.} & \emptyset \rightarrow & \begin{bmatrix} + \text{ son} \\ - \text{ cons} \\ - \text{ syll} \\ + \text{ high} \\ \alpha \text{ back} \end{bmatrix} / [- \text{ cons}] \text{ \_\_\_\_\_\_ } \begin{bmatrix} - \text{ cons} \\ + \text{ high} \\ \alpha \text{ back} \end{bmatrix}
 \end{array}$$

In forms like deyður [dejjUr], from /dej+Ur/, and húðir [húwwIr], from /húw+Ir/, we see that the second element of a diphthong can condition the application of the rule, as well as a vowel. This is why the surrounding segments are specified as [- cons], rather than as [+ syll].

The rules do not apply, however, to break up diphthongs or glide plus vowel sequences within a morpheme. It is therefore necessary to restrict its operation to the position at a morpheme boundary, which will be obtained if we insert a morpheme boundary (+) before the blank in 3.3a, and after the blank in 3.3b. These two rules are obviously related, and just as we collapsed the two rules of glide formation in Icelandic (II.7.5a and II.7.5b) by the mirror image notation, which will be discussed below in chapter IV, so we can collapse the (amended) rules 3.3a and 3.3b above into one schema:

$$3.4 \quad \emptyset \rightarrow \begin{bmatrix} + \text{son} \\ - \text{cons} \\ - \text{syll} \\ + \text{high} \\ \alpha \text{ back} \end{bmatrix} \% \begin{bmatrix} - \text{cons} \\ + \text{high} \\ \alpha \text{ back} \end{bmatrix} + \text{_____} [- \text{cons}]$$

This schema is expanded as the (amended) pair of rules 3.3a and 3.3b, applying in that order. Observe that if they apply conjunctively (that is, if one can apply even after the other has applied), a form like /biw+a/, to be discussed below, will be converted first into /biw+wa/ and then into /biww+wa/. This would, in fact, be incorrect; only one of the two parts of the schema can be allowed to apply. Thus, we see that the rules expanded from 3.4 must be applied disjunctively. Further aspects of this issue will be discussed in section IV.4.

III.4 The Verschärfung. Having discussed some aspects of the vowel system, let us consider the consonantal alternation known as the verschärfung, a name originally given to a sound change in the history of Germanic. This change accounted for correspondances between glides in certain WGmc forms and velar/palatal plus glide clusters in NGmc and EGmc (ddj or ggj = [ʝ] and ggw = [gw]). As Roe pointed out in his recent, highly useful study (Roe 1965), the alternation, which has occasioned as much discussion as virtually any other single fact of Gmc historical phonology except the consonant shifts, has generally been regarded as a peculiarly Gmc phenomenon, with the result that parallels with similar changes in other languages have sometimes been neglected. For instance, Ital maggiore 'greater', from Lat māior shows the same correspondance; in the history of Gujarati the future infix -iya- becomes -iyya- and then -ija-; Prov, Cat, Span, Port guardar, from Gmc wardon 'to ward off'; Span guadanar 'to mow' from Gmc waldanjan, from waida 'meadow'.<sup>6</sup> Examples could

be proliferated from a variety of languages. In each case, the first step in the change is apparently the development of lengthened glides under some circumstances; while the conditions under which these arise remains problematic in some cases, the change itself is of a not uncommon sort, and by no means isolated in Gmc.

Such a change is the basis of a productive rule of verschärfung in modern Faroese. For instance, most verbs that have the clusters ggi (= [ʝ]) and gv (= [gv]) before vocalic endings such as the -a of the infinitive show simply a (diphthongized) long vowel before consonantal endings such as the -r of the 3 sg. pres. of strong verbs:

- 4.1    a. búgva [bɪgva] 'to dwell'; 3 sg. býr [bujr]  
          b. grógva [grægva] 'to grow; 3 sg. grór [grör]  
          c. doyggja [dɔʝa] 'to die'; 3 sg. doyr [dojr]  
          d. spýggja [spUʝa] 'to vomit'; 3 sg. spýr [spujr]

It will be seen that the vowels in the 3 sg. forms of strong verbs have been further affected by an umlaut rule of sorts, which converts back vowels into front ones -- that is, ú/u become ý/y; ó becomes ø; o becomes e; a becomes e, etc. This alternation is completely independent of the verschärfung, and will simply be assumed here.

In 4.1 there is, in addition to the alternation between cluster and glide, an alternation between simple vowels (in the infinitives) and complex vowels (in the 3 sg. forms). There is, furthermore, no dialectal variation in the quality of the nucleus of the reflex of ó in the infinitival forms, though this vowel normally varies between [o], [ø], and [ɛ]:

- 4.2            skógvur [skɛgvUr] 'shoe'  
                  skos [skɛws, sköws, or skows] 'id., gen. sg.'

Furthermore, the nuclear vowel in búgva is [ɪ], rather than [u], the usual

reflex of ú. These various facts make it unappealing to deal with the verschärfung alternation purely in terms of the alternation of consonant with  $\emptyset$ , even with the loss of the second element of a diphthong before such consonant clusters.

Let us see what happens if we represent these forms in the underlying system arrived at above, taking the form without the verschärfung cluster as basic:

- |     |    |          |         |
|-----|----|----------|---------|
| 4.3 | a. | /bú+a/   | /bú+r/  |
|     | b. | /gró+a/  | /gró+r/ |
|     | c. | /dé+a/   | /dé+r/  |
|     | d. | /spí+a/  | /spí+r/ |
|     | e. | /skó+ur/ | /skó+s/ |

After the umlaut rule referred to above has applied, the 3 sg. forms with back vowels (i.e., the second column of 4.3a and 4.3b) will be altered, yielding

- |     |    |         |         |
|-----|----|---------|---------|
| 4.4 | a. | /bú+a/  | /bí+r/  |
|     | b. | /gró+a/ | /grþ+r/ |

Now the diphthongization rule (2.4) will apply to these forms, yielding

- |     |    |           |          |
|-----|----|-----------|----------|
| 4.5 | a. | /biu+a/   | /bui+r/  |
|     | b. | /greo+a/  | /grþ+r/  |
|     | c. | /doe+a/   | /doe+r/  |
|     | d. | /spui+a/  | /spui+r/ |
|     | e. | /skeo+ur/ | /skeo+s/ |

At this point, the first element of each diphthong can receive stress, and the second can be reduced and turned into a glide:

- |     |    |           |          |
|-----|----|-----------|----------|
| 4.6 | a. | /biw+a/   | /buj+r/  |
|     | b. | /grew+a/  | /grþ+r/  |
|     | c. | /doj+a/   | /doj+r/  |
|     | d. | /spuj+a/  | /spuj+r/ |
|     | e. | /skew+ur/ | /skew+s/ |

Now recall the glide formation rule, which was finally formulated as the schema 3.4 above. This rule will apply to these representations, inserting glides after the morpheme boundaries in the forms with vowel endings, but leaving unaffected the forms with consonant endings:

- 4.7    a. /biw+wa/            /buj+r/  
          b. /grew+wa/        /grø+r/  
          c. /doj+ja/           /doj+r/  
          d. /spuj+ja/         /spuj+r/  
          e. /skew+wur/        /skew+s/

At this point the vocalic endings in the forms which are to undergo verschärfung are preceded by geminate glides, while the endings that do not cause verschärfung are preceded by simple glides or simple vowels. We suggested above that it was just this lengthening of glides that gave rise to the corresponding processes in other languages; we could now simply state a rule converting long glides into consonant clusters.

It is possible to relate this rule to another rule in Faroese, however, thus giving at least a partial explanation of its presence in the language. Faroese, like the other West Scandinavian languages, has a rule (or rule schema) which dissimilates the continuance of a sequence of consonantal sonorants. In various WScan dialects, such rules may affect some or all of the clusters nn, rr, ll, nr, rl, etc., converting the first element to a stop (d), thus yielding, e.g., dn, dr, dl, dr, dl, etc. In Faroese, such a rule applies without further restriction to ll, and with the restriction 'after a long vowel or diphthong' to other clusters. This produces alternations such as:

- 4.8    a. seinur [sajnUr] 'late'; seinni [sajdni] 'later'  
          b. morgun [mørgUn] 'morning'; morni [mødni] 'id., dat. sg.'

While the exact structural description for this rule is not clear, the

structural change carried out by it is stated as:

$$4.9 \quad [+ \text{sonorant}] \rightarrow \begin{bmatrix} - \text{sonorant} \\ - \text{continuant} \end{bmatrix} / \begin{bmatrix} \text{---} \\ + \text{cons} \end{bmatrix} \begin{bmatrix} + \text{sonorant} \\ + \text{consonant} \end{bmatrix}$$

(with further restrictions)

Let us assume that this process, whatever its precise statement, was generalized at some point in Faroese to apply not simply to sequences of [+ consonantal] sonorants, but to all sequences of [- syllabic] ones which agree in the feature [ $\pm$  cons]; that is, to glides as well as liquids and nasals. The effect of this on geminate glides, then, will be to turn the first elements of jj and ww into the feature complexes below:

$$4.10 \quad \begin{array}{ll} \text{a.} & \begin{bmatrix} - \text{syll} \\ - \text{son} \\ + \text{cons} \\ - \text{cont} \\ + \text{high} \\ - \text{back} \\ + \text{voi} \\ . \\ . \\ . \end{bmatrix} & \text{b.} & \begin{bmatrix} - \text{syll} \\ - \text{son} \\ + \text{cons} \\ - \text{cont} \\ + \text{high} \\ + \text{back} \\ + \text{voi} \\ . \\ . \\ . \end{bmatrix} \\ & /j/ \rightarrow & & /w/ \rightarrow \end{array}$$

But 4.10a is just the segment [g<sub>1</sub>] (palatalized g), and 4.10b is just [g]. Thus, the result of applying the rule 4.9 to the forms in 4.7 will be to create the forms in 4.11:

- 4.11
- |                            |           |
|----------------------------|-----------|
| a. /big+wa/                | /buj+r/   |
| b. /greg+wa/               | /grø+r/   |
| c. /dog <sub>1</sub> +ja/  | /doj+r/   |
| d. /spug <sub>1</sub> +ja/ | /spuj+r/  |
| e. /skeg+wur/              | /skew+ts/ |

In the position before a front vowel or glide, velars are palatalized to [j] and [č], by a rule which is necessary to produce, e.g.,

- 4.12
- |                      |                                  |                           |
|----------------------|----------------------------------|---------------------------|
| a. stingur [stɪŋgʊr] | 'pain'; stingi [stɪŋji]          | '1d., dat. sg.'           |
| b. liggja [lɪja]     | 'to lie'; underlying /ligg+j+a/; | cf. 3 sg. liggur [lɪggʊr] |



This palatalization rule will affect the first columns of 4.11c and 4.11d, giving

- 4.13 c. /do<sup>Y</sup>+ja/                    /doj+r/  
          d. /spu<sup>Y</sup>+ja/                   /spuj+r/

Now recall the suggestion that the quality of the first elements of ó and ú was produced by a rounding rule applying before rounded vowels (in some dialects, only high rounded ones) or glides. If this rule applies to the forms produced up to this point, we will have an explanation of the lack of rounding in the forms like grógva and búgva, since the segment following the vowel will be a g, not a rounded vowel. In addition, the rule converting w into v (referred to above) must apply to the clusters of g+w in 4.11a, 4.11b, and 4.11e (recall that this rule applies anywhere except after a [+ high] vowel). After these rules have applied, together with the rules for obtaining the phonetic interpretations of vowels, we will obtain the correct phonetic outputs:

- 4.14 a. [bɪgva]                    [bujr]  
       b. [grɛgva]                   [grö:r]  
       c. [dɔ̯ja]                    [dojr]  
       d. [spU̯ja]                   [spujr]  
       e. [skɛgvUr]                [skɛws], [sköws], or [skows]

Thus, all of the rules which are necessary to account for the verschärfung alternation can be shown to be more or less independently motivated in the language if the most natural assumptions are made about the shapes of the underlying forms. In positions other than before vocalic endings, verschärfung, if it occurs, can also be shown to follow from the rest of the grammar. For instance, in a form like hoyggj [hɔ̯j], where the verschärfung appears in final position, it again results from the dissimilation of a geminate glide by rule 4.9; in this case, the first element of the geminate is from the diphthong oy, and the second element represents the effect of a rule which converts final

g to j after a high front vowel or glide (a rule which is independently needed in the grammar). The presence of this underlying g is shown by forms like the genitive [hojks], the g appearing as k before the ending /-s/.

There are other instances in which the verschärfung is unexplained, apparently. Such a form is the word brúgv 'bridge.' In accounting for the second element of the geminate glide that gives rise to the gv cluster here, we can make use of an explanation which is required to explain another set of apparently unmotivated glides. In certain forms, the glide v appears after a non-high vowel and before another non-high vowel, a position in which it could not have arisen regularly by the operation of the glide-epenthesis rule (3.4):

- 4.15 a. æð̥a [æ̥aʋa] 'eider duck'; æð̥u [æ̥aʋu] 'id., oblique cases'  
 b. røð̥a [rø:ʋa] 'speech'; røð̥u [røʋu] 'id., oblique cases'

In these cases, what has evidently happened is that the glide which is present in the oblique case forms has been re-analyzed as part of the root; the stems in 4.15, then, are taken to be /aw/ and /røw/, respectively. If, now, the same reanalysis affects brúgv (in whose other forms, such as the plural brúgv̥ar, a glide is present regularly by 3.4), this stem will then be taken as /brúw/, which will account for the final verschärfung. Outside of the form classes just discussed, there remain only a very few instances of unexplained verschärfung, mostly in the form of consciously preserved archaic doublets.

III.5 Some apparent exceptions. Having now accounted for the forms in which verschärfung occurs, we must now turn to the large class of forms for which, if the underlying form which is apparently motivated by the phonetic forms

were taken as basic, we would expect to find verschärfung on the basis of the rules discussed so far, but do not. In these forms, the conditions appear to exist for the creation of the geminate glides that give rise to the verschärfung clusters, and indeed the forms do contain phonetically long glides, but for some reason these long glides do not dissimilate to form the consonant clusters.

- 5.1    a. víga [vujja] 'dedicate'  
          b. jógur [jüwUr] 'udder'  
          c. síga [sujja] 'lower'  
          d. blíður [blujjUr] 'friendly'  
          e. heiðin [hajjIn] 'heathen'  
          f. týða [tujja] 'translate'

Unless some basis can be found for the non-dissimilation of these glides, we will be left with a large class of forms which must be marked, ad hoc, in the lexicon with an indicator of the non-applicability of the rule. Such an arbitrary sub-categorization of the lexicon is to be avoided if at all possible, of course. A principled explanation might lie either in showing that the long glides have not yet been created at the point rule 4.9 applies, or in showing that some other segment still separates them at this point, a segment which is (presumably) later deleted.

For the forms 5.1a, 5.1b, and 5.1c, an immediate explanation of the latter sort is available. These represent forms with a g between the two vowels in the underlying form. Intervocalic g is everywhere deleted in Faroese; consider some other forms in which these segments are deleted:

- 5.2    a. þga [ð:a] 'increase'  
          b. siga [sija] 'say'  
          c. fagur [fæʌvUr] 'beautiful'

That the underlying forms of 5.1a-c and 5.2 contain a g is demonstrated by the appearance of this segment in other forms of the words, where the g is no longer intervocalic:

- 5.3 a. vǫgdi [vujgdi] 'dedicated'  
 b. júgs [jüwks] 'udder (gen. sg.)'  
 c. sígdi [sujgdi] 'lowered'  
 d. øgdi [ögdi] 'increased'  
 e. sagdi [sagdi] 'said'  
 f. fagran [fagran] 'beautiful (masc. acc. sg.)'

These forms show that we need a rule to delete intervocalic g. If this rule is required to apply after 4.9, it will follow that the forms in 5.1a-c will not show dissimilation, since the segment g will still intervene at the point this rule applies.

The other forms in 5.1 in which dissimilation is absent (5.1d-f) are ones with an ǫ in the orthography, in a position separating the two vowels. The presence in the orthography of such a segment is, of course, no evidence for much of anything, but these are the forms that had a phonetic [ǫ] in this position in Old Norse. Historically, this [ǫ] merged with [ɣ] (the phonetic shape of intervocalic g),<sup>7</sup> and both later disappeared. This disappearance is not, however, of demonstrably greater age than the verschärfung, and in any case, it is by no means implausible that even though they have disappeared from the surface forms, these segments might have left traces in the synchronic grammars of later stages of the language, due to the effect their presence may have had on other phonetic processes. In this case, the effect might remain in the abstract representations of the grammar.

In fact, a segment does show up in the position of the orthographic ǫ in other forms of the same words that fail to show dissimilation. In weak verbs, for example, forms with ǫ show an intrusive d before the d of the preterite, while other forms have only a single d:

- 5.4 a. stríð<sup>ð</sup>a [strujja] 'struggle'; stríddi [strujddi] 'id., pret.'  
 b. týð<sup>ð</sup>a [tujja] 'translate'; týddi [tujddi] 'id., pret.'

Similarly, adjectives of the type that fail to show dissimilation before the ending -ur of the nom. sg. masc. show an extra t before the t of the nom. sg. neut.:<sup>8</sup>

- 5.5 a. blífður [blujjUr] 'friendly'; blítt [bluj<sup>h</sup>t] 'id., neut.'  
 b. fróður [frowUr] 'wise'; frott [frow<sup>h</sup>t] 'id., neut.'  
 c. prúður [pruwUr] 'handsome'; prútt [pruw<sup>h</sup>t] 'id., neut.'

In other such adjectives, the position of the ð is occupied by d before a following n if syncope applies before a vocalic ending:

- 5.6 a. heiðin [hajjIn] 'heathen'; heidnan [hajdnan] 'id., acc.'  
 b. soðin [so:jIn] 'cooked'; sodnan [sodnan] 'id., acc.'

On the basis of these facts, one might well be tempted to set up an underlying dental of some sort in the position occupied by ð in the orthography, and mark it in some way to be deleted later. Other facts, however, argue that the position of the ð is not always occupied by a dental:

- 5.7 a. tráður [troavUr] 'thread'; traðri [træagri] 'id., dat.'  
 b. veður [ve:vUr] 'weather'; veðri [vegrI] 'id., dat.'  
 c. suður [suwUr] 'south'; syðri [sIgrI] 'more southerly'

These forms seem to require a velar in the position of the ð. We can now ask whether there are in fact two such disappearing segments, one a dental and one a velar, or whether they should both be represented in the same way.

On the basis of the facts in 5.7, we can say that if one or the other is to be chosen, it should be the velar. This is because all of the positions in which the segment in question shows up as dental are before another dental, and it would not be implausible to have an underlying velar assimilated to the position of a following dental. There are, in fact, forms in which an overt velar is manifested in similar positions as a dental:

- 5.8 a. draga [dræ:a] 'drag'; drignan [drɪdnan] 'id., masc.  
acc. sg. p.p.'
- b. níga [nujja] 'sink down'; nignan [nɪdnan] 'id., masc.  
acc. sg. p.p.'

Here a g is assimilated to a following n as d, just as would be required for, e.g., 5.6. Only some g's are assimilated, however, and it would be a mistake to represent all of the segments we are concerned with as g. We might, then, represent all of the 'missing segments' as [ɣ], the spirant corresponding to g.

There are other facts, however, that argue that at least some of the 'missing segments' should be represented as dentals. Thus, in the ablaut system, the second class of strong verbs consists of those showing a stem vowel ó or ú. These verbs have these vowels distributed as follows: ó before dentals and before ð, and ú elsewhere:

- 5.9 a. njósa 'sneeze'; njóta 'enjoy'; bjóða 'offer'
- b. krúpa 'creep'; fúka 'rush'; flúgva 'fly'

In order to maintain this generalization, we will have to represent at least these ð's as dentals.

In fact, the only forms for which such a representation would be ill-advised would be the very rare set in 5.7, where there would be no basis for the conversion of an underlying ð into g before r. But the forms in this class are few in number, and there do not appear to be any for which it would not be perfectly possible to represent the segment in the position of orthographic ð as simply g. Thus, these forms have no real bearing on the question of the feature composition of the other ð's, and we should follow the most natural course and represent these latter as dentals. Since t and d already appear in underlying representations, the logical candidate is the corresponding spirant, [ð]. We will then require a rule of spirant deletion:

$$5.10^9 \quad \delta \rightarrow \emptyset$$

Rule 5.10 will eliminate the segment  $\delta$  from phonetic representations, after allowing it to influence other processes. One of its effects will be to be converted to d or t in the environments 5.4, 5.5, and 5.6; another will be to block the dissimilation rule in 5.1d-f. None of the arguments given above as to the feature composition of the segment are in any way relevant to the question of its existence, of course; this is adequately shown by the forms cited.

The forms 5.1d-f, and other similar forms in which dissimilation does not apply, are explained by the presence of either intervocalic g or  $\delta$  in the representations at the point at which rule 4.9 should apply. In order for this explanation to have any force, however, it is necessary to include the requirement that the dissimilation rule not apply after the deletion rule(s). We have already seen that the dissimilation rule must apply after the glide epenthesis rule (3.4), since it is this rule which provides the long glides that are to be dissimilated. Observe, however, that the rule of glide epenthesis must follow the rule of spirant deletion, since the deletion of the spirant results in the insertion of a glide, as shown in the forms of 5.1. We have thus the following ordering relations:

- 5.11
- a. dissimilation precedes spirant deletion (a marked order)
  - b. glide epenthesis precedes dissimilation (an unmarked order)
  - c. spirant deletion precedes glide epenthesis (an unmarked order)

Here of course is another ordering paradox, of the sort which is impossible in a grammar on which a linearization requirement is imposed. On the assumptions adopted here, of course, this solution, which seems the natural and correct one, is perfectly possible; it requires simply the statement, in the grammar of Faroese, of the restriction given as 5.11a above. The other

orderings, being unmarked, follow naturally.

It has been suggested<sup>10</sup> that the paradox of 5.11 can be avoided if, instead of claiming as we do here that the glides in 5.1 are the result of the normal glide insertion rule 3.4, and that this rule can apply after the spirant deletion, we reformulate the spirant deletion rule so as to convert the spirants into glides, rather than deleting them. There are several points against this proposal, however. Firstly, it would require the rule 3.4 to be completely restated, since it is the entire complex of facts captured by that rule that determines which glide is to be inserted. Thus, exactly parallel to 3.4, we would need another rule

$$5.12 \quad \left\{ \begin{array}{c} \delta \\ \gamma \end{array} \right\} \rightarrow \left[ \begin{array}{c} + \text{son} \\ - \text{cons} \\ - \text{syll} \\ + \text{high} \\ \alpha \text{ back} \end{array} \right] \% \left[ \begin{array}{c} - \text{cons} \\ + \text{high} \\ \alpha \text{ back} \end{array} \right] + \text{ \_\_\_\_\_\_ } [- \text{cons}]$$

whose operation exactly duplicates that of 3.4, without being collapsible with it, because at least dissimilation must intervene between the two.

A second objection is that the rules of glide epenthesis and spirant deletion are necessary parts of the grammar in any case. That is, the segment  $\delta$  must appear in underlying representations and does not appear in phonetic representations; in forms like those shown above where it is not converted into something else, and where no glide appears (because both surrounding vowels are [- high]), it must still be deleted. Thus, rule 5.10 must appear anyway; the solution proposed of having a rule like 5.12 would implicitly claim that the glides in the forms in 5.1 are a result of the presence of the spirant, not of the juxtaposition of the vowels. That is, under the most natural application of the rules, the spirants in these forms would be deleted anyway, and the resulting juxtapositions of vowels should lead to



the insertion of glides. Rule 5.12 claims that this is a purely accidental fact. A third objection comes from the fact that at least some of the segments which rule 5.12 must turn into glides are velars. Now there are already rules in the grammar which turn velars into glides, but these are unlike 5.12 in that they are only sensitive to a preceding vowel, and only to a preceding front vowel. Thus, such evidence as there is shows that in Faroese such a process is not as general as 5.12 would claim. Another rule which affects velars, and which is sensitive to both front and back vowels is the velar fronting rule (part of the palatalization process described above); this rule is also sensitive to both preceding and following vowels, but here the following vowel takes precedence, rather than the preceding vowel, as in 3.4 and its siamese twin 5.12. These last two arguments are merely suggestive, of course, of the fact that a process such as that hypothesized in 5.12 would be counter to the 'drift' of Faroese (in something like Sapir's sense), but the other objections, concerning the superfluity and lack of generality of a grammar incorporating 5.12 are sufficient to argue against its existence. In any event, 5.12 has as its only real motivation the avoidance of an ordering condition which, on the assumptions adopted in this work, is perfectly natural.

## Notes

<sup>1</sup> A long syllable is one which is stressed and either open (i.e., followed by only one consonant plus a vowel) or followed by one of the clusters pl, pr, kl, kr, or tr (n.b. not tl). All other syllables are short.

<sup>2</sup> Observe that we have chosen the name [ $\pm$  long] to fill the role of the feature that was called [ $\pm$  tense] in the analysis of Icelandic given in chapter II, and vice versa. This is partly accidental, and partly to emphasize the relation of these two and the arbitrariness of the distinction.

<sup>3</sup> Though this rule may seem simply a fudge, its later role will be seen to have an explanatory function. Such a rule is not unheard of, of course; some dialects of English spoken in the southern and southwestern United States have quite similar processes, whereby the reflex of long u, for instance, may be something like [Iw].

<sup>4</sup> This evidence will not be gone into here in detail; the situation is rather similar to that described in chapter II for Icelandic, though the regularities are even more widespread.

<sup>5</sup> One difficulty might arise if the change of au to [ej] is accepted as very early, as is traditionally assumed. In this case, the split of the two [ej]'s would be hard to explain. However, the [ej] from au is still structurally /au/, and hence distinct from the other [ej]. Further, the only evidence for the early date of au  $\rightarrow$  ej is the palatalization rule, which is adequately explainable in terms of the surface phonetics. For the

argument, cf. Haegstad (1917).

<sup>6</sup> For these examples, cf. Elcock, 1960, (for Romance) and Dave, 1935, (for Gujarati).

<sup>7</sup> It is interesting to note that this strange change of ð to ɣ also occurs in the contiguous Celtic dialects of the Hebrides. There is clearly something common to the two, but the direction of causation is unclear.

<sup>8</sup> The geminate t is realized as preaspirated in many dialects, thus remaining distinct from single t. The character of the preaspiration is the same as that described for Icelandic in chapter II.

<sup>9</sup> It is probably possible to generalize this rule so as to account for the deletion of intervocalic g at the same time; it would be extremely reasonable to have g go to ɣ in this position, in which case the segments to be deleted would be the only voiced spirants in the language. Other apparent voiced spirants (in particular, v) are the result of late rules converting w into v, as described above, and f into v in intervocalic position.

<sup>10</sup> Independently by Arthur Schwartz and Morris Halle (personal communications). Incidentally, the plausibility of such a solution is quite low, since a relation between a dental spirant and a glide is rather far-fetched.

## Chapter IV - Notational conventions

IV.1 Evaluation procedures and notations. Thus far we have not attempted to provide more than brief remarks to clarify the status of the various notational devices that have been employed as part of the phonological theory underlying the work of chapters I - III. We have assumed as given approximately the framework of theoretical assumptions of Chomsky and Halle, 1968. It is the purpose of this chapter to examine, and in some respects to refine that framework.

As has often been pointed out (e.g., in Chomsky, 1967; Chomsky and Halle, 1967; Halle, 1962; and numerous other references), linguistic theory must attempt to provide not only a framework in which the complete range of generalizations about the structures of individual natural languages can be expressed, but also some notion of what a natural language is. As part of this latter project, it is necessary to provide some way of expressing the extent to which a proposed statement does, in fact, capture a possible general property of a possible human language; that is, of the degree of linguistically significant generalization of the rule. To some extent, the use of a phonological feature system for the statement of rules accomplishes this end. Thus, since the features in such a system are defined on the basis of the properties that play a role in language, they admit of a relatively straightforward definition of the naturalness of a class of segments that play a part in a given rule: of two given classes, the one definable in terms of the smaller number of such features is the more general. There are, of course, numerous complications, but this definition makes many choices correctly, and it serves as a useful first approximation. On the basis of it,

we can define at least part of the notion of naturalness of a rule: of two proposed rules, the one statable in terms of more natural classes (i.e., the one whose formulation involves the smaller number of features) is the more natural. Again, numerous problems arise, and some of these problems will be brought up again in chapter VI below, but this will serve as a useful first approximation.

It must be emphasized that this matter is an empirical issue. It is only in terms of an idea of what constitutes linguistically significant generality that we can formulate proposals as to the feature system, etc., that make up the framework in terms of which the rules we are to evaluate by the criteria given above are to be stated; given such a proposed framework (or some significant part of it) we can then proceed to confront it with the vast quantities of empirical data that are available in the languages of the world. The results can only be evaluated with reference to their conformance with our notions of what constitute 'correct descriptions', i.e., descriptions that capture linguistic reality in a significant way. If, for example, some proposal in phonology leads to the circumstance that some rule that the linguist feels to be of a very general nature is highly complex in its formulation while some less general process can be stated simply and elegantly, this must be regarded as counter-evidence to the theoretical proposal on which it is based. The notion that the fundamental empirical basis of claims in linguistics rests ultimately on the linguist's intuition is apparently paradoxical; it is in fact no more so than the claim that any other subject matter is 'empirical'. Large areas of agreement exist among investigators as to the content of a proposed field (that is, what kind of phenomena are to be included in the subject matter of the field); theories are then evaluable in terms of their relative ability to characterize

accurately this range of phenomena. This includes not only the expression of all of the possible phenomena in the field (an essentially taxonomic pursuit, if followed in isolation), but also the delimitation of such phenomena from others. Fortunately, a sufficient community of opinion exists among linguists as to what constitutes a linguistically significant generalization to permit agreement of sorts on what constitutes empirical evidence for or against theoretical proposals.

After these meta-theoretical excesses, we return to the question of the evaluation of the generality of phonological rules. In addition to expressing the naturalness of a set of segments of a language, we want to express the naturalness of sets of strings of segments. For example, if a language has several rules expressing the same change occurring in several different positions, we want to express the extent to which these different positions form a coherent or natural class. That is, given the two sets 1.1a and 1.1b, we want to express the sense in which we feel 1.1b is more general:

1.1	a.	$a \rightarrow b / c \_\_\_ d$	b.	$a \rightarrow b / c \_\_\_ d$
		$a \rightarrow b / e \_\_\_ f$		$a \rightarrow b / e \_\_\_ d$
		$a \rightarrow b / g \_\_\_ h$		$a \rightarrow b / f \_\_\_ d$

In 1.1b, the process ' $a \rightarrow b$ ' applies before d and after any one of c, e, or f. In 1.1a, however, there is no unifying principle to the environments, and this set of processes must accordingly be regarded as less general. To express this sort of generality, we introduce the notational convention of the braces notation: if two or more consecutive strings of symbols  $S_1$  and  $S_2$ .... differ only in that  $S_1$  contains the substring  $s_1$  where  $S_2$  contains  $s_2$ , etc., their complexity is determined by enclosing  $s_1$  and  $s_2$ , ..., in braces, and locating them in one occurrence of the containing string at the point

at which  $S_1$  and  $S_2$  differ. That is, given 1.2a, we say that its complexity is determined from 1.2b:

- 1.2    a.    abcdef  
               abgdef  
               abhdef  
               abstdef
- b.    
$$ab \left\{ \begin{array}{c} c \\ g \\ h \\ st \end{array} \right\} def$$

We say that 1.2b is a schema representing the sequence of strings in 1.2a. We can now apply essentially the same feature counting technique described above to obtain a measure of the relative generality of the sequence 1.2a in terms of the number of features involved in its schematic representation as 1.2b. This notational convention, then, makes the claim that exactly the same sorts of things that make a single rule more natural can contribute to the naturalness of a set of rules, and that further, certain formal resemblances between the component rules of a set (i.e., those captured by the notation) also contribute to generality.

As an example of this notational device, let us consider the well-known process of consonant gradation in Finnish. As a result of this process, the following changes take place at the beginning of a non-initial, short, closed syllable:

- 1.3    a. Geminate stops (i.e., pp, tt, and kk) are simplified to single stops.
- b. Single stops assimilate completely to a preceding liquid or nasal with which they agree in the feature [ $\pm$  coronal] (i.e., mp, nt, rt, lt, ŋk become mm, nn, rr, ll, ŋŋ, resp., while, e.g., k in lk is unaffected).
- c. Except as provided in b, single stops are affected as follows: p becomes v; t becomes d; k becomes j after h, j after a liquid and before a front unrounded vowel, v between two high rounded vowels, and  $\emptyset$  elsewhere.

In a series of papers, McCawley (1964, 1966, 1967) has proposed an analysis of a great deal of Finnish phonology, including the gradation phenomena discussed here. He assumes that the complex set of reflexes of single stops described in 1.3c should be derived by separate rules from an intermediate representation as voiced stops, and that the gradation rule should simply turn p, t, and k into b, d, and g in these positions, the other rules applying later to these to produce the correct outputs. This position is accepted as the starting point of the following discussion.

McCawley proposes that the formulation of the gradation rule as a process of voicing voiceless stops can be extended to include the other cases of 1.3a and 1.3b, as well as 1.3c. Later rules would then operate on these voiced stops, converting them back into voiceless ones in the positions in which gradation does not occur (e.g., in the cluster tk and after s); performing the assimilations required for 1.3b, and deleting the voiced stops after homorganic voiceless ones (with the effect of simplifying geminates -- i.e., pp → pb → p). This treatment allows gradation to be considered a single rule (stop becomes voiced at the beginning of a short, closed, non-initial syllable), but it is not obvious that it is correct. Aside from the fact that the rule restoring the original voiceless stops in non-gradating clusters is an otherwise unnecessary complication, native speakers suggest reasons to believe that there are at least two separate rules involved. In foreign words, the rules which apply to single stops (i.e., 1.3b and 1.3c) are generally not applied, while the simplification of geminates (1.3a) extends to all forms, including loans. In order to state this generalization, it is necessary to recognize a distinction between the gradation processes applying to simple and to geminate stops. Since there is no reason to believe that the geminates go through an intermediate (and highly unusual) stage as



clusters with heterogeneous voicing, the extension of the voicing process to include the degemination cases of gradation as well as the others is suspect.

Consider, then, the following proposed set of rules:

1.4<sup>1</sup>

- a.  $\begin{bmatrix} -\text{cont} \\ -\text{son} \end{bmatrix} \rightarrow [+ \text{voi}] / [+ \text{son}] \text{ --- } \begin{bmatrix} \text{str}\#1 \\ +\text{syll} \end{bmatrix} (j) [-\text{syll}] [-\text{syll}]$
- b.  $\begin{bmatrix} -\text{cont} \\ -\text{son} \end{bmatrix} \rightarrow [+ \text{voi}] / [+ \text{son}] \text{ --- } \begin{bmatrix} \text{str}\#1 \\ +\text{syll} \end{bmatrix} (j) [-\text{syll}] \#$
- c.  $\begin{bmatrix} -\text{cont} \\ -\text{son} \\ \alpha\text{cor} \end{bmatrix} \rightarrow \emptyset / \begin{bmatrix} \text{---} \\ -\text{voi} \end{bmatrix} / \begin{bmatrix} -\text{cont} \\ -\text{son} \\ \alpha\text{cor} \end{bmatrix} \text{ --- } \begin{bmatrix} \text{str}\#1 \\ +\text{syll} \end{bmatrix} (j) [-\text{syll}] [-\text{syll}]$
- d.  $\begin{bmatrix} -\text{cont} \\ -\text{son} \\ \alpha\text{cor} \end{bmatrix} \rightarrow \emptyset / \begin{bmatrix} \text{---} \\ -\text{voi} \end{bmatrix} / \begin{bmatrix} -\text{cont} \\ -\text{son} \\ \alpha\text{cor} \end{bmatrix} \text{ --- } \begin{bmatrix} \text{str}\#1 \\ +\text{syll} \end{bmatrix} (j) [-\text{syll}] \#$
- e.  $\begin{bmatrix} -\text{cont} \\ -\text{son} \\ \alpha\text{cor} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{son} \\ \beta\text{lat} \\ \gamma\text{nas} \end{bmatrix} / \begin{bmatrix} \text{---} \\ +\text{voi} \end{bmatrix} / \begin{bmatrix} -\text{cont} \\ \alpha\text{cor} \\ \gamma\text{nas} \\ \beta\text{lat} \end{bmatrix} \text{ --- } \begin{bmatrix} \text{str}\#1 \\ +\text{syll} \end{bmatrix} (j) [-\text{syll}] [-\text{syll}]$
- f.  $\begin{bmatrix} -\text{cont} \\ -\text{son} \\ \alpha\text{cor} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{son} \\ \beta\text{lat} \\ \gamma\text{nas} \end{bmatrix} / \begin{bmatrix} \text{---} \\ +\text{voi} \end{bmatrix} / \begin{bmatrix} -\text{cont} \\ \alpha\text{cor} \\ \gamma\text{nas} \\ \beta\text{lat} \end{bmatrix} \text{ --- } \begin{bmatrix} \text{str}\#1 \\ +\text{syll} \end{bmatrix} (j) [-\text{syll}] \#$

Obviously, the braces notation can be applied to the ends of the environments, collapsing a with b, c with d, and e with f. The resultant final environment in each case is:

$$1.5 \quad \dots / \text{ --- } \begin{bmatrix} \text{str}\#1 \\ +\text{syll} \end{bmatrix} (j) [-\text{syll}] \left\{ \begin{array}{c} [-\text{syll}] \\ \# \end{array} \right\}$$

Since the three schemata resulting from this first application of the notation all share the environment 1.5, they can be further collapsed by putting braces around the remainders of the rules, obtaining:

1.6

$$\begin{array}{l}
 \text{(i)} \left\{ \begin{array}{l} [-\text{cont}] \\ [-\text{son}] \end{array} \right\} \rightarrow [+ \text{voi}] / [+ \text{son}] \text{ ---} \\
 \text{(ii)} \left\{ \begin{array}{l} [-\text{cont}] \\ [-\text{son}] \\ [\alpha \text{cor}] \end{array} \right\} \rightarrow \emptyset / \left[ \begin{array}{c} \text{---} \\ -\text{voi} \end{array} \right] / \left\{ \begin{array}{l} [-\text{cont}] \\ [-\text{son}] \\ [\alpha \text{cor}] \end{array} \right\} \text{ ---} \\
 \text{(iii)} \left\{ \begin{array}{l} [-\text{cont}] \\ [-\text{son}] \\ [\alpha \text{cor}] \end{array} \right\} \rightarrow \left\{ \begin{array}{l} +\text{son} \\ \gamma \text{ lat} \\ \beta \text{ nas} \end{array} \right\} / [+ \text{voi}] / \left\{ \begin{array}{l} [-\text{cont}] \\ [\alpha \text{cor}] \\ \beta \text{ nas} \\ \gamma \text{ lat} \end{array} \right\} \text{ ---}
 \end{array} \right\} / \text{---} \left[ \begin{array}{l} \text{str} \neq 1 \\ +\text{syll} \end{array} \right] (j) [-\text{syll}] \left\{ \begin{array}{l} [-\text{syll}] \\ \# \end{array} \right\}$$

Introducing a few redundant specifications and taking advantage of the effects of previous rules to eliminate others, we can further simplify this to:

1.7

$$\left\{ \begin{array}{l} [-\text{cont}] \\ [-\text{son}] \\ [\alpha \text{cor}] \end{array} \right\} \rightarrow \left\{ \begin{array}{l} [+ \text{voi}] / [+ \text{son}] \text{ ---} \\ \left\{ \begin{array}{l} \emptyset / \left[ \begin{array}{c} \text{---} \\ -\text{voi} \end{array} \right] \\ \left\{ \begin{array}{l} +\text{son} \\ \gamma \text{ lat} \\ \beta \text{ nas} \end{array} \right\} \end{array} \right\} / \left\{ \begin{array}{l} [-\text{cont}] \\ [\alpha \text{cor}] \\ \beta \text{ nas} \\ \gamma \text{ lat} \end{array} \right\} \text{ ---} \end{array} \right\} / \text{---} \left[ \begin{array}{l} \text{str} \neq 1 \\ +\text{syll} \end{array} \right] (j) [-\text{syll}] \left\{ \begin{array}{l} [-\text{syll}] \\ \# \end{array} \right\}$$

This schema expands (essentially) to 1.6, which in turn expands to 1.4. Where it is necessary in the ensuing discussion to refer to the subparts of this rule, this will be done in terms of the parts of rule 1.6.

In rule 1.7, the common environment (that is, 1.5) simply abbreviates 'at the beginning of a non-initial short closed syllable,' i.e., in gradating position. Part i fills in the [+ voice] feature in single stops (those preceded by a vowel or consonantal sonorant) required by 1.3c, and also as part of the assimilation described in 1.3b. Subrules ii and iii apply only to stops preceded by segments with which they agree in the feature [ $\pm$  coronal]

(i.e., to geminate stops and to the stops that are to be affected by 1.3b). Part ii is further restricted to apply only to voiceless stops; since all stops preceded by sonorants will have been voiced by part i, the only ones that will be affected by this subrule will be the second elements of geminates, which will be simplified. Subrule iii, then, will convert all remaining members of the class (that is, just the stops which have been voiced by part i because they appear after a homorganic consonantal sonorant) into copies of the preceding sonorants. This effects all the changes in 1.3 (assuming later rules to account for 1.3c, as mentioned above) without affecting those clusters not subject to the gradation rule at all.

While the full expansion in 1.4 is quite complex, the schematic representation in 1.7 requires only 14 features for its statement (counting  $\emptyset$  as one, and disregarding the common environment 'at the beginning of a non-initial short closed syllable'). The solution proposed by McCawley (1967) requires 24 features (again, exclusive of the final environment) for the statement of the rules performing the equivalent operation. Thus we see that by the approximate evaluation criterion suggested above, more generality is attributed to this formulation than to McCawley's, and it is thus preferred as part of the grammar of a natural language. Such feature counting arguments cannot be taken too seriously unless they are confirmed by other facts; in this case such added evidence is available. As was seen above, it is necessary to provide some device for stating the exceptionality of loans with respect to subrules 1.3b and 1.3c without thereby entailing exceptionality with respect to 1.3a; in McCawley's framework this is impossible without the addition of another complex minor rule, since he assimilates all of the operations of gradation to a single process of voicing. In this solution, however, it is possible to state that the loan words are exceptional

with respect to subrules i and iii of 1.6, but not to subrule ii;<sup>2</sup> that is, loans are marked [- rule 1.6i] and [-rule 1.6iii]. No words, however, are ever [- rule 1.6ii]. Such additional evidence confirms both the approximate metric of naturalness provided by feature counting and the notational devices (such as the braces notation) that reduce 1.4 to 1.7.

Recall that, although for the purpose of evaluation of the rules in 1.4, they collapse into the schema 1.7, for the purpose of application they remain a set of  $6^3$  ordered rules, applying in the sequence of the subrules of 1.4. That is, the device of the braces notation affects only the evaluation, and not the manner of application of the sequence of rules 1.4.

In fact, the device of abbreviation by braces is not sufficient, even when supplemented by other notational conventions to be discussed below, to fully capture the naturalness of sets of rules. Thus, although all of the subparts of the schema 1.7 are in some sense related in that they are all processes of 'weakening' this fact is not expressed. The set of rules would be no more complex if, instead of [+ son] in subpart iii, we substituted the specification [+ delayed release], and deleted the specifications of laterality and nasality entirely. This would have the consequence of turning stops into affricates after sonorant consonants. Such a set of processes would be highly unnatural, because the weakening processes of parts i and ii would be unrelated to the strengthening performed by the new part iii; however, this fact would not be revealed, and in fact the new formulation would be several features more concise by the evaluation method so far suggested. The coherence and generality of the processes abbreviated by 1.7 is left unexpressed, which means that our proposed evaluation method, while an approximation to a correct one, is not the final answer. Another example of this sort of failure is the following: in many

languages (e.g. Russian), there is a rule which assimilates the voicing of an obstruent to that of a following obstruent. It is often the case that final obstruents are devoiced in such languages as well. There is no way to combine these two statements into a single rule without taking some such ad hoc step as that of Zwicky (1965), who gives the boundary # in Sanskrit the features [+obst, -voice]. Such a step defeats the aim of a phonetic motivation for the features used in phonological description. It is still necessary to express the fact, however, that final devoicing is related in some sense to voicing assimilation; both the positions 'before a voiceless obstruent' and 'finally' are characterized by non-vibration, or unvoicing. As such, the two rules should be evaluated as parts of a single, more general process, but since the devices available to phonological theory so far are all based solely on the formal structure of the rules, there is no way to accomplish this without deforming the feature system. This issue will arise again in chapter VI.

IV.2 Some exceptional types of ordering. The kind of rules that were presented in section I.1 as motivation for the notion of order, and the rules discussed in section IV.1 above, were all cases in which the ordering involved is for one rule to precede another. That is, the input to the rules is operated on by one, and then the result of this process is used as the input to the other, the output being a function of both rules. Such a relationship between rules may be called conjunctive order, after Chomsky (1967). It was pointed out there that examples such as those discussed so far establish the claim that:

2.1 The rules of a grammar must be (partially) ordered.

If all relations between rules were of the conjunctive sort, it might be



The two dialects apparently have the same underlying forms, and differ in that certain optional rules may apply in CC but not in LL. Consider the following pairs of forms:

2.4	LL	CC	
	a. ti:den	teyden	'week'
	b. kruti:	krutey	'cruel'
	c. mi:dlo	meydlo	'soap'
	d. mle:ko	mli:ko	'milk'
	e. ne:st	ni:st	'carry'
	f. dobre:	dobri:	'good'
	g. u:řat	ouřat	'office'
	h. u:doli:	oudoli:	'valley'

Thus, we can describe the relation between LL and CC here by including the rules described in 2.5 in the grammar of CC:

- 2.5
- Add a glide [j] or [w] after tense stressed i:, u: respectively.
  - Exchange tense stressed i: and e:, u: and o:.
  - Lax vowels before glides.

The formulation of subrules 2.5a and 2.5c presents no particular problems, but subrule 2.5b is a case of the sort being discussed here. It must apparently consist of the two parts 2.6a and 2.6b:

- 2.6
- $$\begin{bmatrix} + \text{ high} \\ - \text{ low} \end{bmatrix} \rightarrow [- \text{ high}] / \begin{bmatrix} + \text{ syll} \\ + \text{ str} \\ + \text{ tense} \end{bmatrix}$$
  - $$\begin{bmatrix} - \text{ high} \\ - \text{ low} \end{bmatrix} \rightarrow [+ \text{ high}] / \begin{bmatrix} + \text{ syll} \\ + \text{ str} \\ + \text{ tense} \end{bmatrix}$$

If we abbreviate these two rules by means of the variable  $\alpha$  which ranges over the values + and -, <sup>4</sup> we obtain the schema 2.7:

$$2.7 \quad \begin{bmatrix} \alpha \text{ high} \\ - \text{ low} \end{bmatrix} \longrightarrow [-\alpha \text{ high}] / \begin{bmatrix} + \text{ syll} \\ + \text{ str} \\ + \text{ tense} \end{bmatrix}$$

Clearly the two parts of the expansion of 2.7 (i.e., 2.6) are so related that neither may be allowed to apply to the output of the other. Such a relation between rules, as opposed to the conjunctive ones discussed above, can be called a disjunctive ordering relation. If we make the claim that all sets of rules which are abbreviable by means of the use of a variable are expanded in such a way that their constituent subrules are disjunctively ordered, we have limited this kind of exception to cases where a certain distinguishing formal property is present. If this formal property is present, only one of the pair can apply to any given form; since such sets of rules inherently apply to disjoint sets of representations, the question of which applies first never can arise, and the fact that they apply disjunctively is the only relevant property of the ordering relation between them.

We have, then, discovered a class of counterexamples to 2.2, discovered that examples of this formal sort do exist in natural languages, and suggested that they are characterizable in terms of a formal property (the property of abbreviability under the variables notation). We could thus add this formal property as a codicil to 2.2, and attempt to maintain that, except for cases of this formal sort, 2.2 holds.

It is not sufficient, however, to provide for cases abbreviable by variables as the only counterexamples to 2.2. Consider rules of the form of 2.8:

$$\begin{array}{ll} 2.8 & \text{a. } X \longrightarrow Y / A \text{ \_\_\_\_\_\_ } BC \\ & \text{b. } X \longrightarrow Y / A \text{ \_\_\_\_\_\_ } C \end{array}$$



The use of brackets as suggested in section 1 allows us to collapse these two rules into the schema 2.9:

$$2.9 \quad X \longrightarrow Y / A \text{ --- } \left\{ \begin{array}{c} BC \\ C \end{array} \right\}$$

This is clearly inadequate, however. The strings 'BC' and 'C' have more in common than two other randomly selected strings of the same complexity; in fact, the content of rule 2.8 is that the process 'X  $\longrightarrow$  Y' takes place, not in two distinct environments, but rather in the environment 'before C, whether or not B intervenes.' To capture this, we establish a new notational device for sets of rules related as in 2.8:<sup>5</sup> the element B in such a set is enclosed in parentheses for the purpose of evaluation of the set of rules. That is, 2.8 can be evaluated in terms of the schema 2.10:

$$2.10 \quad X \longrightarrow Y / A \text{ --- } (B)C$$

We define this schema as representing the pair of rules 2.8, applied in that order.

Let us now consider an example of a set of rules related in this way. In Postal (1968; pp. 144-8), a set of rules from Mohawk is presented in support of the claim that ordering relationships have to be defined among the rules of the grammar. Two of these rules are the following:

2.11 a. 'Stress'

$$[+ \text{ syll}] \longrightarrow [+ \text{ stress}] / \text{ --- } [- \text{ syll}]_0 [+ \text{ syll}] [- \text{ syll}]_0 \#$$

b. 'Stress jump'

$$\begin{array}{ccccc} [+ \text{ syll}] & [- \text{ syll}]_0 & \left[ \begin{array}{c} \text{#} \\ +\text{stress} \end{array} \right] & \longrightarrow & \left[ \begin{array}{c} 1 \\ +\text{stress} \end{array} \right] & 2 & \left[ \begin{array}{c} 3 \\ -\text{stress} \end{array} \right] \\ 1 & 2 & 3 & & & & \end{array}$$

The symbol # in 2.11b is an ad hoc cover symbol for a certain class of epenthetic vowels (with quality [a]) which cannot be stressed. Since the

point of rule 2.11b is to skip over such a vowel in determining the location of the stress, and these vowels (being epenthetic) do not exist at all levels of representation, the most obvious way to account for this situation would be to order the rule of epenthesis after the stress rule, so that the latter will operate on representations in which the offending vowels have not yet been inserted. There is not enough data in Postal's presentation to determine whether there is any reason not to adopt this proposal, but since it was not adopted by Postal, especially in a section of a book which is devoted to the issue of ordering, we can assume that there is some reason to assume that these vowels have to be present in the representations to which the stress rule applies. Even on this assumption, however, there is no need for rule 2.11b to appear in the grammar of Mohawk at all. Observe that the two conditions under which stress is assigned are as follows:

- 2.12 a. .../ \_\_\_ [-syll]<sub>o</sub> Δ [-syll]<sub>o</sub> [+syll] [-syll]<sub>o</sub> #  
 b. .../ \_\_\_ [-syll]<sub>o</sub> [+syll] [-syll]<sub>o</sub> #

Since the stress assigned to an Δ by rule 2.11a is always retracted by 2.11b, and seems to play no role, there is no reason not simply to assign stress by the two principles in 2.12 instead. Now the rules 2.12 are related as were 2.8; they can, accordingly, be collapsed into a schema:

- 2.13 [+syll] → [+stress] / \_\_\_ ([-syll]<sub>o</sub> Δ [-syll]<sub>o</sub> [+syll] [-syll]<sub>o</sub>

Observe, however, that if 2.13 (=2.12) is to have the same effect as 2.11, it is necessary to prevent 2.12b from applying to forms to which 2.12a can apply. If both subrules apply, the epenthetic vowel Δ will receive stress, in addition to the vowel preceding it which is supposed to receive the stress. Thus, we see that these rules must apply disjunctively, and that they appear to constitute another class of exceptions to 2.2.

It has been observed in numerous works since the notation of parentheses was first suggested that the rules it abbreviates are always disjunctively ordered, and that furthermore they always apply in the order whereby the longer of the two precedes. In Chomsky (1967) it is proposed that the relation of parentheses - abbreviability is equivalent to this kind of ordering relationship; that is, not only are all rules which are so abbreviable disjunctively ordered, but also all rules which are disjunctively ordered (except those discussed above under the variable notation) are so abbreviable. Examples will be presented below in sections IV.4 and IV.5 that this latter, stronger claim cannot be maintained, but so far as is known the claim that all rules abbreviable by parentheses are disjunctively ordered is strongly supported by the evidence.

We have now established two formally defined classes of exceptions to principle 2.2: sets of rules abbreviable by variables over feature coefficients, and sets of rules abbreviable by parentheses. The peculiar characteristic of both of these classes is that only one of the pair of rules can be allowed to apply at a given place in a given form; that is, that they are disjunctively ordered. The difference between the two kinds of case is that an order of applicability is defined by the notation in the cases abbreviated by parentheses, while no such order exists (or could exist) in the case of rules abbreviated by a variable. Thus, the 'simultaneous' order defined by the variable notation (as this kind of order has sometimes been called) is really just a special case of the other kind of disjunctive order. We are left, therefore, with the following weakened form of principle 2.2:

- 2.14 The rules of a grammar are linearly ordered, except that if two rules are adjacent and are related in one of the formal ways that enforces disjunctive order (i.e., by the variable or parentheses notations), only one

applies to the exclusion of the other.

We shall examine in later sections of this chapter and the next the extent to which this modified principle can be maintained.

IV.3 Iterated application and infinite schemata. So far we have assumed a theory of the manner in which rules are to be applied without making it explicit. In Chomsky and Halle (1968, appendix to chapter 8), a formalism is developed within which a more precise approach to the problem is possible. The procedure developed there is approximately as follows: to apply a rule, the string in question is scanned for positions in which the structural description of the rule is satisfied. Such segments as the rule is to affect are somehow designated, and all of the changes are then carried out at the same time. The rule is not then permitted to re-apply to its own output (unless by some other principle, such as the transformational cycle). Certain sorts of cases exist in language which Chomsky and Halle propose to treat by an extension of this principle; these cases form the subject matter of this section.

The processes in question are those that apply 'all the way across' a long form, such as the following: in Finnish, every odd syllable from the beginning of the word is stressed (i.e., the first, the third, the fifth, etc.).<sup>6</sup> In a case like this, there appears to be no way to state a single rule which will be applicable in all of the required positions; rather, there appear to be a (potentially) unlimited set of such rules, of the sort 'stress the initial', 'stress the third syllable', 'stress the fifth syllable', etc.

Chomsky and Halle approach such processes by introducing a new operator, '( )\*', defined approximately as follows:

- 3.1 The string  $X(Y)*Z$  is a schema representing the potentially infinite set of strings 'XZ', 'XYZ', 'XYYZ', 'XYYYYZ', etc., with each one obtained from the preceding one by the addition of another instance of the substring 'Y'.

Using this definition, we can formulate the (set of) Finnish stress rules as follows:

$$3.2 \quad [+syll] \longrightarrow [+stress] / \#C_o(V_o C_o V_o C_o)^* \text{ ---}$$

The set of rules abbreviated by 3.2 will contain a rule for the first syllable (by 0 substitutions of the enclosed substring), a rule for the third (with one substitution), a rule for the fifth (with two substitutions), etc. The principle of rule application described above is then generalized to schemata of the sort defined by 3.1 by the proviso that a string is scanned for positions in which the structural description of any one of the rules abbreviated by such a schema is satisfied, all at the same time, and all modifications produced by the operation of any one of them are carried out at the same time. Thus, all of the infinite set of rules abbreviated by such a schema are applied simultaneously. This solution seems to provide a fully adequate solution to problems such as the Finnish stress rule just described, and the very similar Southern Paiute stress rule discussed in SPE, but it is not clear that it can be generalized to all cases of a similar nature.

Part of the reason for the success of this device in dealing with the problems of the Finnish and Southern Paiute stress rules is the extreme simplicity of the iterated expression (the expression enclosed in the star brackets) in each. The formalism developed in SPE does not extend beyond the case in which this expression is a simple string (i.e., not a complex schematic expression involving other abbreviatory devices such as parentheses

and braces). Let us consider a more complex case, and examine whether it is possible to extend the device beyond its present limits.

Tübatulabal is a Uto-Aztecan language not too far removed from Southern Paiute (the language discussed in SPE). As described by Voegelin (1936), Tübatulabal is also subject to a principle of alternating stresses, but the principle is rather more complicated than the one described for Southern Paiute. The fundamental principle is that every second mora from the main stress gets a secondary stress. The determination of the location of the main stress appears to be a completely independent fact, determined by a complex range of phonological, morphological, and syntactic factors, and will not be gone into here. The application of secondary stress, however, is determined by the following factors:

- 3.3 a. If the vowel preceding the stress is a long vowel (= two morae), it takes the stress.
- b. If the vowel preceding the stress is a short vowel, the stress goes on the next preceding vowel.
- c. A glottal stop<sup>7</sup> counts the same as a short vowel for the purpose of 3.3b.
- d. For the purpose of applying the stress, any of the following units receive stress on their first element: a short vowel (= one mora); a long vowel (= two morae); a short or long vowel followed by a glottal stop optionally followed by another short or long vowel.

In addition to the above, certain morphemes in the language have inherent stress, and even when these positions do not receive primary stress, they count stressed. Stress proceeds backwards from the position of main stress according to the principles 3.3, except that if a stressed vowel intervenes between the last vowel stressed and the vowel to be affected next, this

latter stress is not applied, but the process of stress assignment goes on from the position of inherent stress. These factors will perhaps become clearer through the examples given in 3.5; the basic principle of stress assignment according to 3.3 is given in 3.4:

$$3.4 \quad [+syll] \rightarrow [+stress] / \text{---} (? ( [+syll] \atop [-stress] )) C_o \left( \left\{ \begin{array}{l} +syll \\ -long \\ -str \end{array} \right\} C_o \left[ \begin{array}{l} +syll \\ +str \end{array} \right] \right)$$

3.5 ĩmbĩwĩbaʔat 'he is wanting to roll string on his thigh'  
2 2 1

yu:udu:yu:udat 'the fruit is mashing'  
2 2 2 1

pĩtĩpĩtĩ:dinat 'he is turning it over repeatedly'  
2 2 1

kuʔudzubil 'the little one'  
2 1 (VʔV counts as one unit)

na:adĩʔi 'the cat (obj)'  
2 2 1 (? counts as a mora)

tĩkapĩganān 'the one who was eating'  
2 2 1

(a:dawĩ:k) tĩkapĩganayĩn '(he saw) the one who was eating'  
2 1 2 2 2 1 (the element /-pĩgana-/ has a fixed stress on the last V)

As we see, the principle 3.4 seems to be applicable before stresses assigned by prior applications of 3.4, as well as before stresses which are either lexical or assigned by the primary stress rules. This iterative application of a rule like 3.4 is of course contrary to the principle of rule application suggested above. Is it possible to avoid this conclusion, and formulate the rule as a single infinite schema? The following might be suggested:

3.6

$$\begin{array}{c}
 [+syll] \rightarrow [+stress] / \text{---}((\left[ \begin{smallmatrix} +syll \\ -str \end{smallmatrix} \right]))C_o \left( \left\{ \begin{smallmatrix} ? \\ \left[ \begin{smallmatrix} +syll \\ -long \\ -str \end{smallmatrix} \right] \end{smallmatrix} \right\} \right) C_o VC_o * \\
 \hline
 \begin{array}{c}
 \text{A} \\
 (\left[ \begin{smallmatrix} +syll \\ -str \end{smallmatrix} \right])C_o \left( \left\{ \begin{smallmatrix} +syll \\ -long \\ -str \end{smallmatrix} \right\} \right) C_o \left[ \begin{smallmatrix} +syll \\ +str \end{smallmatrix} \right] \\
 \hline
 \text{B}
 \end{array}
 \end{array}$$

Observe that the A part of the rule simply restates the environment given in the B part, allowing it to appear any number of times between the vowel affected and the stressed vowel whose presence causes the rule to apply. Observe that the principle of simultaneous application has already forced us to state the rather complicated conditions of 3.4 twice, while the rule 3.4 allows them to be stated only once if it is allowed to apply iteratively. As such, the principle of simultaneous application forces us to miss the fact that the A and B parts of 3.6 are essentially the same -- that is, that the B part is not simply an arbitrary string of symbols of a given complexity. Still, it is a worthwhile project to attempt to limit the number of primitives of the theory by giving a single principle of rule application and prohibiting iterative rules. It has already been shown that not all rules can be allowed to apply iteratively -- hence, it would be necessary to mark every rule as to whether it applies in this way or not. It might be worth the added complexity in the formulation of rule 3.6 to avoid this.

Observe, however, that the expansion of 3.6 into a set of simultaneously applicable rules will be a difficult task. Since the string enclosed in iteration brackets is complex, it represents subrules which must have an ordering defined among themselves by the principles of expansion of the braces and parentheses notations. Since these subrules are ordered,



there are going to be problems in the formalism of applying them simultaneously. Now the extension of the principle of application of the subrules of an infinite schema to include sub-schemata involving parentheses need present no particular problems. Since the principle of expansion of parentheses specifies that the subrules are to be disjunctively ordered, it would be possible simply to pick the first applicable member of the expansion as the subrule which represents this 'degree' of iterative expansion. This would give a unique representative of the disjunctive schema for each degree of expansion of the iterative schema, and these 'designated' rules could be applied simultaneously. If the expression to be iterated contains braces, however, no such natural extension is available. By the definition of the braces notation, rules related in this way must be applied in a conjunctive sequence. Thus, if any rule exists which contains braces in an expression to be iterated, more than one subrule from each degree of iteration will have to apply, and they must apply in sequence. Thus, some of the expansion will have to apply to the outputs of other parts of the expansion, contrary to the principle of simultaneous application.

In fact, in the example from Tübatulabal just discussed, there is such a use of braces, but the fact of the conjunctive application of such subrules makes no difference to the output. This is an idiosyncrasy of this case however; in fact, it is in general difficult to find cases where the conjunctiveness of the expansion is important (though the Finnish case just discussed is such a case, as is the English stress rule as given in SPE). Where it makes a difference, however, the evidence supports the assertion that the braces notation enforces conjunctive application; there is no reason whatsoever to believe that this is only true of the uses of braces that are not involved in infinite schematic representations. Thus, even in the absence

of a suitable crucial case, it is clear that either the principle of simultaneous application must be revised, or the generality of the principle of conjunctive expansion of braces must be abandoned. In light of the evidence that already exists for the latter, and the fact that the former leads to problems of formulation to which there is a ready alternative (i.e., the device of allowing some rules, such as 3.4, to apply to their own output), it seems natural to abandon the infinite schema approach, at least for some problems. The fact that rules must then be marked as to whether or not they apply to their own output is not really a serious problem; there is no a priori reason to assume that such a mark is more costly than the mark which is necessary to indicate that a rule contains an infinite schema. We therefore conclude that rules must be admitted which apply to their own output; whether it is possible then to eliminate entirely the device of infinite schemata is unclear. There is certainly a close relation between the two devices, and if it is necessary to admit both in linguistic theory, there may be generalizations missed thereby. The necessity of the iterated application of rules seems clear enough, however, and a case will be presented below in Chapter V where it admits a simple solution to an otherwise complex problem.

IV.4 Neighborhood rules. The abbreviatory devices discussed thus far in this chapter have been well established parts of linguistic theory almost since the beginning of concern with explicit formal representations of grammars. More recently, however, a new device has been suggested in a paper of Bach's (1968; an earlier version of this was read at an LSA meeting in December, 1966). This is the so-called neighborhood, or mirror-image rule, which makes the claim that if a phonological process applies in



opposite sequence; i.e., whether 4.2 is expanded as 4.1 or as 4.4:

- 4.4    a.  $X \rightarrow Y / \text{ \_\_\_\_\_\_ } A$   
           b.  $X \rightarrow Y / A \text{ \_\_\_\_\_\_ }$

This choice, which must be made universally as part of the definition of the notation 4.2, is forced upon Bach, as is the choice of the location of the blank in the expanded subrules as a consequence of the formal properties of the notation he has selected. It may seem superfluous to point this out, but it is useful to note that, while a notational convention can be dreamed up to abbreviate literally anything, the particular choice of notations, once made, can have empirical consequences in the form of choices forced on its definitions which can be tested. Where you save does make a difference. Bach's device makes the very specific claim that the only sets of rules that exist in natural languages whose environments are symmetrical will have both the shape and the order of 4.3. This very strong claim would be an extremely interesting result if supported, and is completely open to testing.

Let us consider some simple cases of neighborhood rules. In Yuma (cf. Halpern, 1946), the basic vowel a is completely assimilated to the quality of another vowel from which it is separated only by a glottal stop. This fact could be expressed by the pair of rules 4.5:

- 4.5
- a.  $\begin{bmatrix} + \text{ syll} \\ + \text{ low} \end{bmatrix} \rightarrow \begin{bmatrix} \alpha \text{ high} \\ \beta \text{ low} \\ \gamma \text{ back} \\ \delta \text{ round} \end{bmatrix} / \begin{bmatrix} + \text{ syll} \\ \alpha \text{ high} \\ \beta \text{ low} \\ \gamma \text{ back} \\ \delta \text{ round} \end{bmatrix} ? \text{ \_\_\_\_\_\_ }$
- b.  $\begin{bmatrix} + \text{ syll} \\ + \text{ low} \end{bmatrix} \rightarrow \begin{bmatrix} \alpha \text{ high} \\ \beta \text{ low} \\ \gamma \text{ back} \\ \delta \text{ round} \end{bmatrix} / \text{ \_\_\_\_\_\_ } ? \begin{bmatrix} + \text{ syll} \\ \alpha \text{ high} \\ \beta \text{ low} \\ \gamma \text{ back} \\ \delta \text{ round} \end{bmatrix}$

These rules will apply to forms such as those in 4.6, giving the outputs shown:

- 4.6 a. /viʔavak/ → [viʔivak] 'here I am'  
 b. /caʔu:lʲ/ → [cuʔu:lʲ] 'arrow-point'

The two rules in 4.5 can be abbreviated (correctly) by Bach's convention to form the single schema:

$$4.7 \quad \begin{bmatrix} + \text{syll} \\ + \text{low} \end{bmatrix} \rightarrow \begin{bmatrix} \alpha \text{ high} \\ \beta \text{ low} \\ \gamma \text{ back} \\ \delta \text{ round} \end{bmatrix} / \begin{bmatrix} + \text{syll} \\ \alpha \text{ high} \\ \beta \text{ low} \\ \gamma \text{ back} \\ \delta \text{ round} \end{bmatrix} ?$$

The fact that the affected segment appears at the end of the string is consistent with Bach's claim. Since there are apparently no strings in the language of the form  $V_1?a?V_2$  where  $V_1 \neq V_2$ , it is not possible to tell in which order the rules should expand, since both orders will yield the correct results for the forms that exist. Thus, this example is consistent with Bach's hypothesis on all counts.

The type of example on which Bach based his proposal was the following: in English, and many other languages, the frontness of velars is determined by the frontness of the adjacent vowels. Thus, after front vowels, velars are front and after back vowels back, unless they are also followed by a vowel, in which case the frontness of the following vowel prevails. This situation can be described as follows:

- 4.8 a.  $\begin{bmatrix} - \text{son} \\ + \text{high} \end{bmatrix} \rightarrow [\alpha \text{ back}] / \begin{bmatrix} + \text{syll} \\ \alpha \text{ back} \end{bmatrix} \text{ —}$   
 b.  $\begin{bmatrix} - \text{son} \\ + \text{high} \end{bmatrix} \rightarrow [\alpha \text{ back}] / \text{ — } \begin{bmatrix} + \text{syll} \\ \alpha \text{ back} \end{bmatrix}$

These rules are abbreviated as:

$$4.9 \quad \begin{bmatrix} - \text{son} \\ + \text{high} \end{bmatrix} \rightarrow [\alpha \text{ back}] / \begin{bmatrix} + \text{syll} \\ \alpha \text{ back} \end{bmatrix}$$

In this case the order of expansion does make a difference in the effect of the rules; since it is that of 4.8, which is formally equivalent to 4.1, this example is consistent with Bach's proposal.

Consider the rule of glide-formation in Icelandic, however, which was discussed as rule II.7.5 above. This rule consisted of the two parts

4.10 (= II.7.5)

- a. [- stress]  $\rightarrow$  [- syll] / \_\_\_\_ [+ syll]
- b. [- stress]  $\rightarrow$  [- syll] / [+ syll] \_\_\_\_

The rules must apply in this order, for a sequence of unstressed vowels such as /...i+u.../ must become [...ju...], rather than [...iw...]. This set of rules is formally equivalent to 4.4, not to 4.1; hence, we must conclude there are neighborhood rules in natural languages with both forms, and Bach's notation is in need of revision. Observe, in passing, that the rule of Icelandic glide formation is consistent with the spirit of Bach's notation; this claims that the segments following the affected segment take precedence over those preceding it in determining the application of the rule. It is only the formal nature of the operation performed by the rule which requires the order of expansion to be different in the two cases. It is an interesting property of phonological rules in general that this assymetry is quite pronounced in all types of rules; while there are of course many exceptions, it is still much more common for a change to be effected by following segments than by preceding ones. Whatever the status of this generalization, it is left completely unexpressed in current phonological theory.

That the formal difference in the operations performed by, e.g., rules 4.9 and 4.10 affects the order of expansion follows from the assumption that neighborhood rules are conjunctively ordered, an assumption that will be examined in detail below; for now, we can note that if the rules were

disjunctively ordered, and applied in the opposite orders, the correct results would also be obtained for 4.9, but not for 4.10. This is due to the fact that rules 4.9 feed one another (in the sense of chapter I), while rules 4.10 bleed one another. If two rules stand in this latter relationship to one another, we could speak of them as 'intrinsically disjunctively ordered.'

To examine the other aspect of Bach's claim, that all neighborhood rules are of the form  $X \rightarrow Y / A \text{ \_\_\_\_\_\_ } B$  (and the mirror image of this) where B is null, recall the rule of glide insertion in Faroese. It will be recalled that this rule (III.3.3) involved environments on both sides of the blank:

4.11 (= III.3.3)

$$\begin{array}{lcl}
 \text{a. } \emptyset \rightarrow & \begin{bmatrix} + \text{ son} \\ - \text{ cons} \\ - \text{ syll} \\ + \text{ high} \\ \alpha \text{ back} \end{bmatrix} & / \begin{bmatrix} - \text{ cons} \\ + \text{ high} \\ \alpha \text{ back} \end{bmatrix} + \text{ \_\_\_\_\_\_ } [- \text{ cons}] \\
 \\
 \text{b. } \emptyset \rightarrow & \begin{bmatrix} + \text{ son} \\ - \text{ cons} \\ - \text{ syll} \\ + \text{ high} \\ \alpha \text{ back} \end{bmatrix} & / [- \text{ cons}] \text{ \_\_\_\_\_\_ } + \begin{bmatrix} - \text{ cons} \\ + \text{ high} \\ \alpha \text{ back} \end{bmatrix}
 \end{array}$$

These two rules are clearly related by the neighborhood convention, since the environments are symmetric, but since they contain an environment on each side of the segment affected (in this case, this segment is  $\emptyset$ ), they are not collapsible by Bach's notation.

We have seen, then, that both of the claims made by Bach's choice of notation are false: rules related by the symmetry condition are not confined to those in which the environment on one side of the affected segment in the subrules is null, as is shown by the Faroese case, and (at least if they are to be ordered conjunctively) it is not possible to make the choice of 4.1 versus 4.4 on a universal basis, as is shown by the contrast between the

English and the Icelandic rules. In light of these objections, it is necessary to formulate a new notation for neighborhood rules, to replace that of 4.2;

let us define the schema 4.12

$$4.12 \quad X \longrightarrow Y \% A \text{ \_\_\_ } B \quad (\text{where } A = a_1 a_2 \dots a_n, \text{ and } B = b_1 b_2 \dots b_m)$$

as an abbreviation for the set of rules 4.13:

$$4.13 \quad \begin{array}{l} \text{a. } X \longrightarrow Y / A \text{ \_\_\_ } B \\ \text{b. } X \longrightarrow Y / \bar{B} \text{ \_\_\_ } \bar{A} \quad (\text{where } \bar{A} = a_n a_{n-1} \dots a_1, \text{ and } \bar{B} = b_m b_{m-1} \dots b_1) \end{array}$$

On the basis of this notation, it would not be possible to specify a preferred order of expansion (that is, to choose universally between the order 4.13 and the opposite order), since A and B are taken to be any strings, and hence, in general, not formally distinguishable. Thus, this notation allows either order of expansion to be specified as part of the formulation of the rule: if the order opposite to 4.13 were desired, the rule could have been formulated as 4.14:

$$4.14 \quad X \longrightarrow Y \% \bar{B} \text{ \_\_\_ } \bar{A}$$

Using this new notation, we can abbreviate 4.11 as 4.15:

$$4.15 \quad (= \text{III.3.4})$$

$$\emptyset \longrightarrow \begin{bmatrix} + \text{ son} \\ - \text{ cons} \\ - \text{ syll} \\ + \text{ high} \\ \alpha \text{ back} \end{bmatrix} \% \begin{bmatrix} - \text{ cons} \\ + \text{ high} \\ \alpha \text{ back} \end{bmatrix} + \text{ \_\_\_\_\_\_ } [- \text{ cons}]$$

We have not yet settled the question of whether the subrules expanded from a neighborhood schema should be ordered conjunctively or disjunctively. In none of the cases which we have discussed so far has it been necessary



to enforce conjunctive order; in each case, the operation of each of the subrules is independent of the operation of the other, and while in most cases it would be equally possible to formulate the rules as conjunctive, the general trend is apparently toward disjunctive orders (frequently intrinsic ones). The Faroese case alone seemed to require that the relation of disjunctive order be enforced by the notation, since it is required (as was seen above in section III.3), and not implicit.

Let us consider what kind of case would be required to make a conclusive decision between conjunctive and disjunctive order for such rules. Such a critical example would have the following form: the rule  $R: X \rightarrow Y$  applies to a set of segments  $S_x$ , mapping them onto elements of the set of segments  $S_y$ . Some of the members of  $S_y$  are also members of  $S_x$ ; call the intersection of  $S_x$  and  $S_y$  the set  $S^*$ . For at least one member of  $S_x$ , say  $\underline{s}$ , its correspondent under  $R$  in  $S_y$  is distinct from it and a member of  $S^*$ ; call this element  $\underline{s}'$ . Since  $\underline{s}'$  is a member of  $S^*$ , it is a member of  $S_x$ ; hence it has a correspondent under  $R$  in  $S_y$ , say  $\underline{s}''$ , distinct from  $\underline{s}'$ . Now given a string  $\underline{AsB}$  such that both subrules of  $R$  can apply to  $\underline{s}$  in the environment  $A \text{ \_\_\_\_\_\_ } B$ , it will be possible to determine whether the rule  $R$  is expanded into conjunctive or disjunctive subrules by examination of the reflex of  $\underline{s}$  in  $A \text{ \_\_\_\_\_\_ } B$ . If the reflex of  $\underline{s}$  is  $\underline{s}''$ , the rules must have been conjunctive, with  $\underline{s} \rightarrow \underline{s}' \rightarrow \underline{s}''$ . On the other hand, if the reflex of  $\underline{s}$  is  $\underline{s}'$ , the rules must have been disjunctive, since only one evidently applied; this rule sent  $\underline{s} \rightarrow \underline{s}'$ . Had the rules been conjunctive, the other would have applied as well, sending  $\underline{s}' \rightarrow \underline{s}''$ , since both were, ex hypothesi, applicable in  $A \text{ \_\_\_\_\_\_ } B$ .

As may be imagined, such criterial cases are extremely difficult to find. Most instances of neighborhood rules in natural languages are of the simple sort exemplified by Bach's rules, and are susceptible of either sort

of formulation. A few cases seem to be of the formal structure just outlined, however, and these will now be discussed.

First, it will be noted that the rule of Faroese glide insertion has the structure discussed above. In the position between two high vowels, both parts of the rule could apply. The distinction between s' and s'' of the above discussion as correspondents of s (=∅) is that between one and two inserted glides; as was noted before, only one glide is inserted. It is thus necessary to expand 4.15 disjunctively.

Another example of a criterial case for this discussion is found in Old Breton. Breton, like the other Celtic languages, has undergone mutation processes applying to consonants in certain environments, which have left synchronic reflexes in the grammars of the particular languages. In most cases, the effects of the mutations have simply been incorporated as changes in the lexical items of the language (this has been the case when the conditions for mutation, together with the segment affected, are found entirely within the boundaries of a single form). Alternatively, the mutation may have as reflex a morphologically conditioned alternation. While these alternations were once phonologically predictable, this is no longer the case for most of the modern languages. In modern Breton, for instance (cf. Hardie, 1948), the so-called soft mutation<sup>8</sup> applies to the initial consonant of feminine singular or masculine plural nouns after the definite article; to adjectives following feminine nouns; and after a heterogeneous class of other words and particles. In each case, it is possible to reconstruct a phonological basis for the change in the etymology of the mutation-causing form, but it is impossible to motivate these explanations in a synchronic grammar. In some cases, the mutation has been extended from the original (phonologically defined) class of mutation-causing forms

to other forms that are morphologically but not phonologically related. In Cornish, this process evidently went so far in the last years the language was spoken (presumably because it was increasingly a second language, and hence less well known) that the 'soft' mutation took place in nouns after virtually anything. In Old Breton, however, the basis for the mutations is still transparently phonological. According to Fleuriot (1964, section 79, p. 209):

La lénition processus phonétique (= soft mutation, *sra*) s'est achevée en Brittonique vers la fin du *v<sup>e</sup>* siècle. *p*, *t*, *k*, *b*, *d*, *g*, *m* ont donné *b*, *d*, *g*, *v*, *ð*, *ɣ*, *ɥ* dans les positions suivantes:

a) entre deux voyelles

b) entre une voyelle et une sonante *r*, *l*, *m*, *n*

In the position between a vowel and a sonorant (= case b), the process applies regardless of whether the order is vowel - sonorant or sonorant - vowel. For example, *\*art+petoc* gives *art+bedog*, and *\*are+tren* gives *are+dren* (which later becomes *ardren*). If we recall that vowels are [+sonorant], as well as *r*, *l*, *m*, and *n*, we see that we can formulate this change as a neighborhood rule, applying between a vowel and a sonorant in either order:

$$4.16 \quad \left[ \begin{array}{c} - \text{continuant} \\ < + \text{voice} > \end{array} \right] \rightarrow \left[ \begin{array}{c} < + \text{continuant} > \\ + \text{voice} \end{array} \right] \% [+ \text{syll}] \text{ \_\_\_\_ } [+ \text{son}]$$

It might be argued that it is not necessary to formulate this process as a neighborhood rule, but that it could instead be formulated as a single rule applying between two [+son] segments. This is not the case, however; apparently, clusters where both flanking sonorants are [-syll], such as *ltr*, *rpr*, etc., are maintained unchanged.<sup>9</sup> It thus appears essential that one of the sonorants be [+syllabic], but irrelevant which.

Now observe that the rule applies to voiceless stops, making them

voiced, and that voiced stops are in turn subject to the rule, becoming spirants. Thus, this is a criterial case of the sort we are seeking. If we can find forms in which, e.g., p is subject to both parts of the schema 4.16, we will have evidence for either conjunctive or disjunctive expansion, depending on whether the result is b (= s'), or v (= s''). In fact, such a case arises in every instance of intervocalic p. Since vowels are sonorants, the sequence VpV can be analyzed either as [+syll] p [+son] or as [+son] p [+syll]. Intervocalic p in fact becomes b, and not v, so only one of these analyses can have resulted in an application of the rule. Thus, the two parts of 4.16 must be expanded disjunctively, just as were the parts of the Faroese rule. If the mutation rule is expressed as one rule like 4.13, this must surely be a neighborhood rule, and this neighborhood rule cannot represent a conjunctive schema. That it is in fact one rule is supported by the fact that in the modern language, where the conditions for the applicability of the corresponding change are rather complex morphological ones, they are the same for all of the segments affected.

The examples of neighborhood rule schemata discussed so far all seem to point to such rules' having the property of disjunctive order, just as rules abbreviated by the parentheses notation do. If this conclusion is born out, we might then add another codicil to principle 2.14 above, to assert that the formal relation of symmetric environments also implies exceptionality with respect to ordering.

But in fact the evidence is by no means unequivocal on the side of disjunctive order. In seeking criterial cases (that is, cases having the formal structure described above), we encounter not only cases that argue for disjunctive ordering, but also cases that argue for conjunctive ordering. The first such case to be discussed is really representative of a large class

of similar examples, found in a wide variety of the languages of the world. Perhaps the most frequently encountered type of neighborhood rule is the case in which some feature of the articulation of a consonant is transferred to an adjacent vowel as a non-distinctive secondary feature; e.g., when vowels are nasalized next to a nasal consonant, pharyngealized next to a pharyngeal, etc. In general, such processes consist in the marking of the vowel for the relevant (binary) feature; that is, a vowel may become, say, [+ nasal] in the environment % \_\_\_\_ [+ nasal]. Such a rule does not, of course, provide any evidence relevant to the question at issue here, since this kind of rule does not have the required form. In terms of that discussion, the elements of  $S_y$  onto which the elements of  $S_x$  are mapped by R all are mapped onto themselves by another application of R. That is, once a vowel gets to be [+ nasal], that's it.

Consider the following exception to this general property of such rules, however. According to Keller (1959), in (Mayan) Chontal,

"Vowels contiguous to glottalized consonants, glottal stop, or the allophone of the phoneme /b/ which has glottal quality<sup>10</sup> may be laryngealized. The laryngealization is more pronounced between two such consonants than when contiguous to just one."

An example is the form ?ahpək'əb, where all three vowels are (non-distinctively) laryngealized, but the feature is said to be more prominent on the last, which occurs between (glottalized) k' and the syllable final allophone of b.

The possibility that more than one degree of a given feature must be present in the output at least of the lowest level of phonetic rules of a language has of course long been recognized,<sup>11</sup> and the two degrees of laryngealization reported by Keller for Chontal are clearly to be accounted for in terms of two different numeric values for this feature (whose presence in

underlying representations is presumably only classificatory, i.e., binary, but whose specification in the most detailed level of phonetic representation is in terms of a many-valued, perhaps continuous scale). Let us assume that such a scale for this feature runs from 0 (which would correspond to the classificatory value [- laryngeal]) to, say, 10. Let us arbitrarily locate the first and second vowels of ʔahpək'əb at the phonetic value of [4 laryngeal], and the third vowel at [8 laryngeal]. We could then describe the facts under consideration by the following sequence of phonetic interpretation rules:

4.17<sup>12</sup>

- a.  $\begin{bmatrix} + \text{syll} \\ x \text{ laryngeal} \end{bmatrix} \longrightarrow [x+4 \text{ laryngeal}] / [+ \text{glot cl}] \text{ \_\_\_\_\_\_}$
- b.  $\begin{bmatrix} + \text{syll} \\ x \text{ laryngeal} \end{bmatrix} \longrightarrow [x+4 \text{ laryngeal}] / \text{ \_\_\_\_\_\_ } [+ \text{glot cl}]$

(where x is a variable ranging over numeric values used in much the same way as the variable a, which ranges over the classificatory values + and - .)

It is not possible to determine a correct ordering for these two rules, but it is necessary for them to apply conjunctively if they are to be used to predict the value of [8 laryngeal] for the final vowel of ʔahpək'əb.

The pair of rules in 4.17 are obviously susceptible of abbreviation by the neighborhood convention, giving 4.18:

4.18  $\begin{bmatrix} + \text{syll} \\ x \text{ lary} \end{bmatrix} \longrightarrow [x+4 \text{ lary}] \% [+ \text{glot cl}] \text{ \_\_\_\_\_\_}$

It should be obvious that this case is of the form discussed above, with a representative value for s' being the same vowel with the value [4 laryngeal], and the correspondent of s', s'', being the same vowel with the feature [8 laryngeal]. Since the reflex of s under two possible applications of R (that is, the value of a vowel between two glottalized segments) is s'', not

s', we must conclude that this constitutes evidence for the conjunctive expansion of 4.18.

Similar examples can be found in many languages for the feature [ $\pm$  nasal], etc. Notice that this kind of case could never provide conclusive evidence that neighborhood expansions are disjunctive, for in such a case (i.e., where the numeric value of the feature in question is the same whether the segment is adjacent to one or two of the segments in question) the facts could be analyzed as was suggested at the beginning of this discussion; that is, by a single rule assigning a classificatory feature, whose numeric interpretation comes later and is constant. Since the same classificatory value would be introduced by both rules of the schema, the numeric identity would be explained without invoking a principle of disjunctive ordering.<sup>13</sup>

Another example of an apparent criterial case comes from Old Irish (cf. Thurneysen, 1961). The case in point concerns the treatment of unstressed vowels in closed syllables. Consonants in Old Irish can apparently be of three types: 'neutral', 'i-quality', and 'u-quality'. Let us assume that these latter two kinds have the features [ $\pm$  high] and [ $\pm$  back] (with [ $\pm$  back] corresponding to 'u-quality' and [ $-$  back] corresponding to 'i-quality') in addition to their other consonant features, while 'neutral' consonants are [ $-$  high]. The 'u-quality' consonants would also be [ $\pm$  round]. Disregarding the effects of i-quality consonants, we have the following vowels in the following positions: between two neutral consonants, a; between a neutral and a u-quality, o; between a u-quality and a neutral, o; and between two u-quality consonants, u. Orthographic citations would be unilluminating, since the orthography does not generally mark any distinction of consonant quality, but a complete discussion will be found in section 102 of Thurneysen

(1961).

It seems that this is what is going on here: unstressed vowels are basically a. Adjacent to a u-quality consonant, these vowels are subject to a raising and backing effect. That is, a becomes o adjacent to a u-quality consonant on either side, and u between two of them. Since the process of raising a to o involves a change of the feature [ $\pm$  low], while the change from o to u involves a change of the feature [ $\pm$  high], there is no very good way to formulate this rule so as to express the unity of the changes; one possible (not fully satisfactory) rule would be:

$$4.19 \quad \begin{bmatrix} + \text{syll} \\ - \text{stress} \\ \langle \text{low} \rangle \end{bmatrix} \rightarrow \begin{bmatrix} - \text{low} \\ \langle + \text{high} \rangle \\ + \text{round} \end{bmatrix} \% \begin{bmatrix} + \text{cons} \\ + \text{high} \\ + \text{back} \\ + \text{round} \end{bmatrix} \text{ —}$$

Regardless of the formulation of the rule, it is clear that it has the correct formal character: s is a; s' is o, and s'' is u. Since the reflex of s under two applications of R is u, the rules abbreviated by 4.19 (or whatever is substituted for it) must be applied conjunctively.

We have seen here that in the cases in which it is possible to tell for sure whether a neighborhood rule should be expanded into conjunctively or disjunctively ordered subrules, there are cases supporting both hypotheses. Of the cases discussed above, two were disjunctive (the Faroese and Old Breton cases) while two were conjunctive (the Chontal and the Old Irish cases). The rest were capable of being treated in either way. Since both sorts of cases were well established, it appears that it is simply not possible to arrive at a universally valid generalization about the manner of organization of such schemata. This does not, of course, affect the generalizations that rules abbreviable by the variables or parentheses notations are disjunctively ordered, while those abbreviable under the braces



notation are conjunctively ordered. It simply shows that the property of disjunctive ordering, while always true of parentheses-abbreviable sets of rules, is not equivalent to the formal property of parentheses abbreviability. That is, there are some sets of rules that are disjunctively ordered, but not parentheses abbreviable, but not vice versa. We have seen that this residual disjunctive ordering is not equatable with neighborhood-abbreviability either (since some schemata of this sort are conjunctive), and unless some other formal property can be discovered which exactly characterizes the set of neighborhood rules that are disjunctive, it will be necessary to mark at least all neighborhood schemata for the property of disjunctive versus conjunctive order. The hope of finding such a property seems vanishingly small. In the next section of this chapter, we will see that there are other cases that support the notion that disjunctive ordering is a property of the relation between rules which is not completely reducible to any other formal relationship between them, though of course there are still generalizations to be made, such as those made about the devices of braces, variables, and parentheses. The fact that these generalizations do not exhaust the range of possibilities does not make them any less important.

IV.5 Disjunctive order and the status of notations. As has been noted in the sections above, the use of notational conventions in linguistic theory is not motivated by any aesthetic concern, a desire to conserve printer's ink in presentation, or some a priori sense of simplicity. It would, of course, be possible to construct any number of strange and ad hoc devices for the formulation of particular processes which conformed to any or all of these principles, and perhaps were even usable in a wide variety of languages. An explicit theory of linguistic structure uses notational conventions to make

empirical hypotheses about what sorts of relatedness among rules constitute genuine generalizations, corresponding to linguistic realities, and what sorts of relations are simply fortuitous formal similarities. They make potentially testable claims about what sorts of relations between rules imply that the rules so related are in a sense parts of the same process; as such, evidence bearing on this latter issue can be evidence that the notational conventions capture important aspects of language. The notational conventions are an aspect of the evaluation procedure that expresses the degree of naturalness of a set of rules; therefore, the questions of what rules are related in what ways, what sets of rules are natural, and what kinds of relationship obtain among related rules are all aspects of the same question.

To date, all of the notational devices that have been proposed in generative phonological theory have operated on the strictly formal structure of the rules involved. This is, indeed, implicit in the very notion of notational convention, as a part of the evaluation function. It is also related to the assumption that the complexity of rules can be made to reflect directly their naturalness, by, for instance, counting the features required for their formulation under some particular collapsing operation. The adequacy of this approach to the naturalness of individual rules will form part of the subject of chapter VI; our aim here is to mention the evidence that can be brought to bear on the issue of the reality of the notational conventions themselves.

It is important to separate the role which the sort of formal relationship between rules a notational convention captures plays in the evaluation of grammars from that which it plays in determining aspects of the operation of the grammar, such as the prediction of ordering relationships. We have

seen above that it is possible to maintain that if two rules are formally related by either the parentheses or variables notation, they are applied disjunctively; however, these notations have different consequences for the evaluation of grammars. The braces notation, which has an evaluatory function similar to that of the parentheses notation, implies that the rules so related are conjunctively applied. The neighborhood notation, on the other hand, while playing a role in evaluation somewhat similar to that of the variable notation, seems to leave undetermined the issue of conjunctive versus disjunctive application, which seems to be an ad hoc property of each such set of rules. From this, we see that certain order-determining functions may be performed by more than one notation, while certain notations may admit of more than one type of ordering. This establishes the independence of the two aspects of a notational function.

So far we have said little about whether rules bearing no particular formal similarity to one another are conjunctively or disjunctively ordered, and it has been assumed that they are conjunctive. There exists, however, some evidence that this is not always the case. It is to such an example that we turn now. In Middle English, as in Modern English, there was a rule which shortened the (stressed) first vowel of a trisyllabic word.<sup>14</sup> For example, a bisyllabic word like sori [sɔ:ri] which has a long vowel in the first syllable (from Old English sārig), has a trisyllabic plural form sorie [sɔriə] with a shortened vowel in the first syllable. This rule can be written:

#### 5.1 (Tri-syllabic shortening)

$$\left[ \begin{array}{l} + \text{syll} \\ + \text{stress} \end{array} \right] \rightarrow [- \text{long}] / \text{--- CVC}_0 \text{V}$$



allowed to follow rule 5.1 (at least), for if it did there would be no reason for the short vowels in the initials of the trisyllabic forms. Observe in addition that it is not possible simply to apply the rules in the sequence 5.3 - 5.1; although this would yield the correct values for the feature [ $\pm$  long] in the examples 5.4, the derived vowel quality would be incorrect. Sample derivations of the trisyllabic forms in 5.4 are given in 5.5 under each of the possible ordering assumptions:

5.5	<u>iveles</u>	<u>sikerli</u>	<u>sumeres</u>
underlying:	/ivel+es/	/siker+li/	/sumert+es/
5.1	/ivel+es/	/siker+li/	/sumert+es/
5.3	/ēvel+es/	/sēker+li/	/sōmert+es/
5.3	/ēvel+es/	/sēker+li/	/sōmert+es/
5.1	/evel+es/	/seker+li/	/somert+es/

output: \*[ēveles] \*[eveles] \*[sēkerli] \*[sekerli] \*[sōmeres] \*[someres]

All of these outputs are incorrect. There would thus appear to be no possible order in which the rules 5.1 and 5.3 can apply. Observe that it would not be possible to mark these roots as 'exceptional' with respect to 5.3 (whose quality change is the principal offense of the incorrect outputs on 5.5), since the rule must apply in the derivation of bisyllabic forms of the same words -- e.g., ēvel, sēker, sōmer. In fact, the class of exceptions would be exactly 'forms to which 5.1 also applies.' Surely there would be a generalization going begging here. A notational variant of this solution (with exception features) would involve a reformulation of 5.1 so that it also introduces the exception feature [- Rule 5.3]. Both of these solutions are simply other ways of saying that the rules 5.1 and 5.3 apply disjunctively. We thus apparently have a case of two rules which are not abbreviable by any

of the notations which we have shown to involve disjunctive ordering (parentheses, variables or neighborhood schemata), but which are nevertheless disjunctively ordered.

If we accept this conclusion, it would seem to follow that the property of disjunctiveness must be stated ad hoc for every pair of rules in the grammar, at least unless it can be predicted from a notational similarity. But this is surely the wrong conclusion; among the thousands of rules that have been proposed for the many languages that have been examined in the course of the development of generative phonology, no other pair has ever been presented as an example of disjunctive ordering unless a notational similarity obtained. From this we would surely like to conclude that rules that are unrelated are conjunctive.

An alternative to stating disjunctiveness ad hoc for every pair of rules in every language would be extending the principle of one of the notational conventions to include the case just discussed. We can note that it is not altogether unreasonable to do so; the environments of rules 5.1 and 5.3 are related by the parentheses notation, since one is a substring of the other:

$$5.6 \quad \dots / \text{---} CV(C_0V)$$

Furthermore, the rule corresponding to the longer of the two environments takes precedence, as is asserted to be universally the case with parenthetical schemata. Note also that while the structural changes of the two rules are not identical, as they would have to be to permit the parentheses notation per se to be used, they have a very definite functional relation: both primarily serve to specify the value of the feature [ $\pm$  long]. We are confronted with the following situation: two rules specifying the value of the same feature, applying in environments such that the cases satisfying

one analysis are a subset of the cases satisfying the other, and applying disjunctively with the longer environment taking precedence. This is the same relationship between rules that is captured by the device of the parentheses notation, but without the required formal similarity of operations. It would not seem unreasonable, then, to extend the function of this notation to include cases such as that discussed here, much as it is necessary to extend the function of the braces notation to cover cases such as the subrules of the Finnish gradation rule discussed in section IV.1 above.<sup>15</sup> It might be observed that this extended principle of disjunctive order is quite similar to that employed by Panini (cf. Staal, 1968), which also does not require formal identity of operation, but only that the cases to which the one rule applies be a subset of the class of cases to which the other applies. Since the structural description of a rule is really a specification of an equivalence class of forms in the language, if one SD is a substring of another, the class of forms corresponding to the longer is a subset of the class of forms corresponding to the shorter. Thus, the principle of disjunctive ordering is essentially Panini's principle of the precedence of the more specific rule.

The fact that the relation of disjunctive expansion is consistently correlated with the relation of (extended) parentheses notation is definite evidence that this latter captures an empirically real aspect of linguistic structure. Although we have indicated for both the braces notation (in section IV.1) and the parentheses notation (above) that the linguistic principles they represent are not exactly statable in formal terms, the fact that we can obtain evidence about the kinds of extension that must be made to make them fully descriptive of the facts, and that these extensions appear fairly natural, is an indication that they constitute close approximations

to real and important aspects of the operation of human languages.

Another sort of evidence for the reality of the notational conventions was presented in section II.2 above. There it was noted that the actual class of forms to explain which speakers of Icelandic are required to construct an umlaut rule is such that either a formulation with parentheses or an infinite schema would be possible. These differ, however, in that the parentheses schema would have to apply conjunctively rather than disjunctively, as we have seen to be the case with other disjunctive schemata, if the correct results are to be obtained. Since the two formulations make different predictions about the treatment of forms with more than two 'umlautable' a-vowels, it is possible to determine which formulation more nearly corresponds to the speakers' internal grammar. While a large scale study of this topic has not been attempted, it appears that speaker intuition favors the infinite schema representation. Since this prediction about the treatment of non-occurring forms would be unnecessary if the parentheses notation did not require disjunctiveness, the result that the formulation of the rule in the speakers' grammars is the one avoiding the requirement of disjunctiveness (and thereby making predictions beyond the range of experience on which the grammar is based) is itself strong evidence for the reality of the relation between the parentheses schema and disjunctive ordering.

Still another class of evidence bearing on the question of the realities captured by notational conventions comes from the facts concerning historical change pointed out by Kiparsky (1968b). He noted that rules which a notational convention collapses ought to behave as some sort of a unit in historical change; for example, new rules should be inserted into a grammar in such a way as not to break up such a unit, the rules should participate as a block



in changes such as reorderings, simplifications, etc. If such conditions are always maintained, it would be strong evidence in confirmation of the reality of the notational conventions.

Kiparsky's principal example of this is as follows: in Old English, vowels were laxed under (at least) the following two conditions: before a cluster of three or more consonants; and before two consonants followed by two more syllables. These conditions can be formulated as the two environments for the laxing rule in 5.7 below:

- 5.7 a.  $V \rightarrow [-\text{long}] / \_\_\_\text{CCC}$   
 b.  $V \rightarrow [-\text{long}] / \_\_\_\text{CCVC}_oV$

These obviously collapse to the single schema:

$$5.8 \quad V \rightarrow [-\text{long}] / \_\_\_\text{CC} \left\{ \begin{array}{c} C \\ \text{VC}_oV \end{array} \right\}$$

In Middle English, 5.7a is simplified to apply before any two consonants, and 5.7b to apply before one consonant followed by two syllables. These result in 5.9a and 5.9b, which again collapse, giving the schema 5.9c:

- 5.9 a.  $V \rightarrow [-\text{long}] / \_\_\_\text{CC}$   
 b.  $V \rightarrow [-\text{long}] / \_\_\_\text{CVC}_oV$   
 c.  $V \rightarrow [-\text{long}] / \_\_\_\text{C} \left\{ \begin{array}{c} C \\ \text{VC}_oV \end{array} \right\}$

Obviously, the change has not simply affected the two separate parts of 5.7, but rather the whole schema 5.8, simplifying it to the schema 5.9c. The occurrence of such changes is a demonstration of the reality of such schemata as 5.8 as units in historical change.

It should be pointed out that there is a problem with this example. In Middle English, there is another rule which rounds a:

$$5.10 \quad \begin{bmatrix} + \text{syll} \\ + \text{low} \\ + \text{back} \\ + \text{long} \end{bmatrix} \rightarrow [+ \text{round}]$$

Now at least in some dialect areas,<sup>16</sup> this rule applied before rule 5.9b above; thus, from OE hāligdæg, we get first hōliday, and then holiday. Rule 5.9a, however, seems always to have preceded 5.10; thus, the genitive of hālig, which involves a vowel syncope resulting in a consonant cluster after the original long vowel, was halges. Thus, the rules seem to have applied in the order 5.9a - 5.10 - 5.9b. Such an order would imply that the sound change consisting in the introduction of rule 5.10 into the grammar involved the insertion of a rule between two rules collapsed by a schematic notation. This would seem to prevent their being collapsed after this point, which would constitute counterevidence to the reality of this device.

Consider, however, the fact that the two rules 5.9a and 5.9b never apply to the same form. Thus, on the assumptions about ordering adopted in this work (as outlined in chapter I), their ordering relationships with respect to other rules are completely independent of one another. Since there is no requirement that rules be arranged in a single linearly ordered list, and that only adjacent rules on this list can be collapsed, there is no reason not to collapse rules 5.9a and 5.9b into the single schema 5.9c, despite the fact that they bear different ordering relations with respect to 5.10. These ordering assumptions, then, combine with the historical evidence to support the reality of the braces notation. Other evidence for the reality of the braces notation, on the other hand, supports the assumptions about ordering made in this work, for they make it possible to avoid a potential counter-argument to the reality of this principle. The issue of the empirical

support for the ordering assumptions adopted here, however, will be the subject of chapter V.

## Notes

<sup>1</sup> The various notational devices used in formulating this rule will be examined in detail later; for the purposes of this discussion, it is sufficient to note that the appearance of (j) in these subrules actually represents the abbreviation of a pair of disjunctively ordered rules. The elements  $\underline{\alpha}$ ,  $\underline{\beta}$ ,  $\underline{\gamma}$ , etc., represent variables ranging over the values + and -, as will be explained below. See also note 6 for the use of  $\text{str}\#1$ .

<sup>2</sup> Note that it was possible to eliminate the feature [- son] from the segment in the environment preceding the dash in collapsing this rule because, if the preceding segment were [+ son], the segment represented by the dash would have become [+ voi]. This segment is specified as [- voi], however, which is sufficient (in light of the effects of subpart i) to eliminate the possibility of the preceding segments being other than [- son]. In loan words, however, whose grammar does not include subparts i and iii, this redundancy is not present, and hence the feature [- son] cannot be eliminated.

<sup>3</sup> This really represents a set of 44 rules, if the full implications of the presence of parentheses and variables are realized. These complications play no particular role here, however, and can be disregarded.

<sup>4</sup> To say a variable 'ranges over' the values + and - is to say that its use abbreviates two expressions, in one of which + is substituted for the variable, and in the other of which - is so substituted. The algebra of these elements is as follows:  $--=+$ ,  $-+=-$ .

<sup>5</sup> The position of the distinct substrings in the containing string is irrelevant, of course; the braces can potentially occur in any part of the rule.

<sup>6</sup> The initial stress is primary, and all following stresses are secondary; thus, in the discussion in part 1 above, the use of 'str#1' means anywhere but in the first syllable.

<sup>7</sup> Only occurrences of glottal stop that are present in underlying forms count for this purpose; thus, there are later rules that insert glottal stops, but these apply after the stress rule, and hence these stops are not counted. This comment applies, e.g., to the final glottal stop in the form ĩmbĩŋwibaʔat in 3.5 below.

<sup>8</sup> This mutation has been traditionally called by various names in the various daughter languages, such as 'spirant mutation', etc., according to the modern reflexes of the results.

<sup>9</sup> For details, cf. Fleuriot (1964), sections 39-48.

<sup>10</sup> This allophone occurs in syllable final position.

<sup>11</sup> Reference to the use of numeric values as feature coefficients will be found at various points in SPE.

<sup>12</sup> The use of the feature names 'laryngeal' and 'glottal closure' should not be taken too seriously here; recent work by L. Maran (unpublished)

has shown promise of illuminating the structure of the feature system for these problems, but the structure of the rule here is independent of the particular features chosen.

<sup>13</sup> It has been suggested that low-level rules such as the one just discussed should not be offered as evidence for the structure of phonological theory, since the facts they express are notoriously dependent on the vagaries of individual phoneticians. Published descriptions, however, sometimes indicate that the phonetic facts concerning some low-level feature are similar to those discussed in connection with [ $\pm$  laryngeal] in Chontal, and sometimes that only one degree of (non-distinctive) laryngealization, nasality, etc., is present. As such, it would seem to be a language particular fact that one language contains a rule like 4.18, while another contains [ $\pm$  syll]  $\rightarrow$  [ $\pm$  laryngealized] % [ $\pm$  glot cl] \_\_\_\_\_. Since these language particular facts are presumably to be expressed as rules, there is no reason to exclude them from an evidential function. One cannot but deplore the non-uniformity of phonetic description that raises this indeterminacy, however.

<sup>14</sup> For details, cf. Wright (1928) and Moore (1951).

<sup>15</sup> Of course, some of the generalizations involved in the Finnish gradation rule discussed in section 1 were adequately handled by the unextended braces notation -- that is why the example was brought up there. This reference is rather to the remarks at the end of that section concerning the relation between 'softening' processes, and the distinction between these and 'hardening' processes.

<sup>16</sup> For details, cf. Brunner (1963) and the references of fn. 13.

## Chapter V - The Principles of Rule Ordering

V.1 Introduction. In chapter I, it was observed that rules must be applied in at least a partial order if the correct generalizations about phonological structure are to be stated in a unified way. We examined there a set of assumptions about the ordering relation that must, therefore, be defined on phonological rules; these assumptions were described as the requirement of linearizability. They impose a set of conditions on the rules of the phonology, such that if they are met it would be possible to give a single linear order for the rules, such that applying them in that order would yield correct results for every form. We suggested that the requirement of linearizability might be incorrect, and proposed an alternative theory of the phonological ordering relation, to which we gave the name local ordering.

In chapter IV, we have examined some of the characteristics of the ordering relation, and have discovered classes of cases for which a strict principle of linear order is violated. These are the cases of disjunctively ordered rules, cases which are more or less well defined on the basis of the formal properties of the rules. These cases, while they violate a strict requirement of linear order by providing cases where the application of a given rule entails the skipping over of another rule in the order, do not violate the requirement of linearizability. As such, they are not directly relevant to the establishment of the theory of local order.

In the course of discussion of the vowel systems of Icelandic and Faroese in chapters II and III, however, we encountered several cases in which the most natural solution seemed inconsistent with the requirement of linearizability, but which were adequately (and hopefully, elegantly) handled within the theory of local ordering. These cases, of which the Icelandic



u-umlaut and the Faroese glide epenthesis rules are perhaps the best examples, provide corresponding support for this theory. There are still a number of problems of definition involved in the concept of local order, however; most notably in the matter of the definition of 'unmarked order'. Furthermore, these few cases from two closely related languages are hardly sufficient to establish as universal a principle as radically different from current assumptions as this. Therefore, this chapter is devoted to exploring the notions of local order and marked/unmarked order, and to providing some other examples of problems whose treatment within this framework appears more natural than any possible within the framework of linear ordering.

V.2 Menomini. Bever (1967) has recently discussed the theoretical background of the description of Menomini given in Bloomfield's work (1939; 1962). In the process, he has reformulated much of this description in terms of generative phonology. Among other results, he has concluded that the underlying vowel system of the language consists of the four elements (plus length):

2.1	e(:)	o(:)
	æ(:)	a(:)

He further concludes that it is not necessary to assume the existence of any basic semi-vowels in the underlying representations of Menomini, for all instances of phonetic [j] and [w] can be derived predictably from underlying instances of /e/ and /o/. His evidence for these claims will not be repeated here in detail.<sup>1</sup> According to Bever, the rule producing the glides is (reformulated in the feature system in use here):

$$2.2 \quad \begin{bmatrix} + \text{ syll} \\ - \text{ long} \\ - \text{ low} \end{bmatrix} \rightarrow \begin{bmatrix} - \text{ syll} \\ + \text{ high} \end{bmatrix} / ([+ \text{ syll}]) \text{ --- } \left\{ \begin{array}{l} [+ \text{ syll}] \\ + \end{array} \right\}$$

This rule, in turn, expands to the following four subrules:

$$2.3 \quad \begin{bmatrix} + \text{ syll} \\ - \text{ long} \\ - \text{ low} \end{bmatrix} \rightarrow \begin{bmatrix} - \text{ syll} \\ + \text{ high} \end{bmatrix} / \left\{ \begin{array}{l} \text{A) } [+ \text{ syll}] \text{ --- } [+ \text{ syll}] \\ \text{B) } [+ \text{ syll}] \text{ --- } + \\ \text{C) } \text{ --- } [+ \text{ syll}] \\ \text{D) } \text{ --- } + \end{array} \right\}$$

Cases A and B of this rule are, in accordance with the conventions discussed in chapter IV, disjunctively ordered with respect to cases C and D. These rules apply to produce derivations like these:

2.4	base form:	/oe:keoam/	/oeskoo/	/kone/
	2.3A	/oe:kewam/	"	"
	2.3B	"	/oeskow/	"
	2.3C	/we:kewam/	/weskow	"
	2.3D	"	"	/kony/
	output:	[we:kewam]	[weskow]	[kony]

Bever discusses in some detail the necessity of ordering the subrules of 2.3 in this way, and the incorrect consequences of any other ordering.

Another rule discussed by Bever deals with sequences of vowels not altered by 2.3. A short vowel drops adjacent to a long vowel; and the second of two short vowels drops. This can be formulated as the following pair of rules:

$$2.5 \quad \begin{array}{ll} \text{a. } \begin{bmatrix} + \text{ syll} \\ - \text{ long} \end{bmatrix} \rightarrow \emptyset / [+ \text{ syll}] + \text{ ---} \\ \text{b. } \begin{bmatrix} + \text{ syll} \\ - \text{ long} \end{bmatrix} \rightarrow \emptyset / \text{ ---} + [+ \text{ syll}] . \end{array}$$

These two rules can be collapsed, of course, by the neighborhood convention discussed in section IV.4, giving:

$$2.6 \quad \begin{bmatrix} + \text{ syll} \\ - \text{ long} \end{bmatrix} \rightarrow \emptyset \% [+ \text{ syll}] + \text{ ---}$$

A further wrinkle is that one of a sequence of a: vowels drops (say the second). Bever attempts to combine this process with rule 2.6, with indifferent success. We can state it as 2.7, without taking a stand on any possible relation it might bear to 2.6:

$$2.7 \quad \begin{bmatrix} + \text{ syll} \\ + \text{ low} \\ + \text{ back} \end{bmatrix} \rightarrow \emptyset / \begin{bmatrix} + \text{ syll} \\ + \text{ low} \\ + \text{ back} \end{bmatrix} \text{ — } \text{ — }$$

If 2.7 follows 2.6, there is no need to specify that the two vowels involved are long; if either of them were short, it would have been deleted by the operation of 2.6. Observe that 2.6 must follow 2.2, since if one of the members of a vowel sequence is e or o, its conversion to a glide by 2.2 takes precedence over its deletion by 2.6. The reader can satisfy himself of the incorrectness of the results of applying 2.6 to the forms in 2.4 before 2.2 applies.

Any remaining sequences of vowels (only sequences of long vowels remain after the operation of 2.6 and 2.7) are broken up by the insertion of a y-glide. This could be formulated as in 2.8:

$$2.8 \quad \emptyset \rightarrow \begin{bmatrix} - \text{ syll} \\ - \text{ back} \\ - \text{ low} \\ + \text{ high} \\ + \text{ son} \end{bmatrix} / [+ \text{ syll}] \text{ — } [+ \text{ syll}]$$

There is another, obviously related, rule which inserts an e to break up a sequence of non-syllabic plus true consonant. This rule could be formulated as in 2.9:

$$2.9 \quad \emptyset \rightarrow \begin{bmatrix} + \text{ syll} \\ - \text{ back} \\ - \text{ high} \\ - \text{ low} \\ + \text{ son} \end{bmatrix} / [- \text{ syll}] \text{ — } \begin{bmatrix} - \text{ syll} \\ - \text{ son} \end{bmatrix}$$

These rules can be collapsed into one, as shown here, since vowels are

[+ sonorant]:

$$2.10 \quad \emptyset \rightarrow \begin{bmatrix} \alpha \text{ syll} \\ - \text{ back} \\ - \alpha \text{ high} \\ - \text{ low} \\ + \text{ son} \end{bmatrix} \quad / \quad [\alpha \text{ syll}] \quad \text{---} \quad \begin{bmatrix} \alpha \text{ syll} \\ \alpha \text{ son} \end{bmatrix}$$

Notice that while this rule correctly states the facts, it fails to capture the fact that the segments e and y in Menomini are very closely related: as was pointed out above, it is possible to derive every instance of surface y from underlying e except, apparently, those produced by the operation of rule 2.10. Yet rule 2.10 would be no simpler, would express no greater generalization, if by judicious use of alphas, it inserted some other vowel, such as u, unrelated to y, to break up the prohibited clusters. Further, 2.10 misses the fact that all of the instances of y that are produced by it are in positions where underlying e would have been turned into y if it had appeared there. If rule 2.10 were reformulated to insert e in all cases, and rule 2.2 were permitted to apply to its output, these generalizations could be obtained. The new insertion rule would look like this:

$$2.11 \quad \emptyset \rightarrow \begin{bmatrix} + \text{ syll} \\ - \text{ high} \\ - \text{ back} \\ - \text{ low} \\ + \text{ son} \end{bmatrix} \quad / \quad [\alpha \text{ syll}] \quad \text{---} \quad \begin{bmatrix} \alpha \text{ syll} \\ \alpha \text{ son} \end{bmatrix}$$

Recall that we have thus far established the following ordering relationships: 2.2 precedes 2.6; 2.6 precedes 2.7; and both 2.6 and 2.7 precede 2.11. None of these orderings are unmarked. So far, we have not established any order between 2.2 and 2.11, if we assume the orderings established are not transitive. In that case, if the epenthesis process is formulated as 2.11, the order 2.11 - 2.2 is a feeding order, and hence unmarked. No undesirable consequences can result from allowing 2.2 to apply

after 2.11 as well as before a rule that precedes 2.11 (i.e., 2.6 and 2.7), since the only instance of vowels to which it could apply in the former environment are those created by 2.11, all others having been removed by 2.6 or previous application of 2.2. Consider the following derivations, adapted from Bever (1967), assuming the vowel assimilation processes that apply after all rules considered here:

2.12 underlying form: /ase æ: + æn + am/  
           by 2.2:       asy æ: + æn + am  
           by 2.6:       asy æ: + n + am  
           by assim:     [asi: nam]       (= output)

underlying form: /ase æ: + æ: n + æm æo/  
           by 2.2:       asy æ: + æ: n + æm æw  
           by 2.11:     asy æ: + e æ: n + æm æw  
           by 2.2:       asy æ: + y æ: n + æm æw  
           by assim:     [asi: y æ: n æm æw]   (= output)

underlying form: /naka: + a: pe æ: + naso æ + æo/  
           by 2.2:       naka: + a: py æ: + nasw æ + æw  
           by 2.6:       naka: + a: py æ: + nasw æ + w  
           by 2.7:       naka: + py æ: + nasw æ + w  
           by assim:     [naka: pi: nasow]   (= output)

underlying form: /po: n + m/  
           by 2.11       [po: nem]       (= output)

underlying form: /kea: so + t/  
           by 2.2:       kya: sw + t  
           by 2.11:     kya: sw + et  
           by assim:     [kya: sot]       (= output)

This treatment seems more natural than that which treats the two epentheses as separate processes, and the essence of the improvement is the relaxation of the requirement of transitivity of ordering. While the process described

here is hardly sufficient by itself to motivate the relaxation of the requirement of transitivity, taken together with the other examples in this chapter and in chapters II and III, the increased generality of the relation between e and y in Menomini provides some added evidence for such a step.

V.3 Kasem. In SPE, Chomsky and Halle discussed a portion of the phonology of Kasem, an African language. Their discussion was based on an article by Callow (1965). A further article of Callow's (1968) provides additional data which support the analysis proposed by Chomsky and Halle, though emphasizing the fact that the rules are only applicable to a portion of the total language.

A large number of nouns in Kasem share the suffixes /+a/ 'singular' and /+i/ 'plural'. These are the class C nouns of Callow's discussion, such as:

- 3.1 a. bakada 'boy'; pl. bakadi
- b. sada 'grass mat'; pl. sadi
- c. fala 'white man'; pl. fali
- d. bakala 'shoulder'; pl. bakali

If a form would otherwise end in two identical vowels, one of these is truncated. This rule applies only in some classes of words, but class C nouns are among them:

- 3.2 a. kambia 'cooking pot'; pl. kambi, from /kambi+i/
- b. pia 'yam'; pl. pi from /pi+i/

Some forms do not show this truncation; one such (daa 'stick') is cited by Chomsky and Halle, who suggest that it might be simply a transcription error. The further data of Callow (1968), however, indicate that non-truncation is not isolated to this particular form, but rather that a large class of forms are systematically exceptional with respect to the rule.

However these exceptions are to be specified (i.e., regardless of whether the individual forms should be marked as exceptions, or the rule itself altered so that the exceptions are built into its environment), the truncation rule will look like 3.3:

$$3.3 \quad \begin{bmatrix} + \text{syll} \\ \alpha \text{ high} \\ \beta \text{ back} \end{bmatrix} \rightarrow \emptyset / \text{---} \begin{bmatrix} + \text{syll} \\ \alpha \text{ high} \\ \beta \text{ back} \end{bmatrix}$$

Other rules discussed by Chomsky and Halle include velar elision, which applies before the suffix /+i/ to delete any velar (e.g., buga 'river'; pl. bwi), and glide formation, which converts a high vowel into a corresponding glide before a vowel. Glide formation does not apply before the suffix /+a/; it may well be possible to restrict its application to the position before a [- low] vowel. While sequences of glide plus a do occur, they are apparently confined to positions entirely within the morpheme, and the segment represented by a glide never alternates with a vowel in these forms. It would seem natural, then, to represent these segments as underlying glides, rather than as vowels; in this case, the glide formation rule can be stated simply as:

$$3.4 \quad [+ \text{high}] \rightarrow [- \text{syll}] / \text{---} \begin{bmatrix} + \text{syll} \\ - \text{low} \end{bmatrix}$$

Another process in the language, similar to the phenomenon of guṇa in Sanskrit, is the contraction of sequences of low vowel plus high vowel into a single mid vowel; that is, /ai/ = [e]; /au/ = [o]:

$$3.5 \quad \begin{bmatrix} + \text{syll} \\ + \text{low} \end{bmatrix}_1 \begin{bmatrix} + \text{syll} \\ + \text{high} \end{bmatrix}_2 \rightarrow \emptyset \begin{bmatrix} 2 \\ - \text{high} \end{bmatrix}$$

This rule plays a part in derivations like these:

- 3.6 a. laja 'song'; pl. le from /la+i/; from /laŋ+i/  
 b. naga 'let'; pl. ne from /na+i/; from /naŋ+i/

in which the velar of the stem disappears by the velar elision rule, and the /a+i/ sequence is then reduced to [e] by rule 3.5.

With these rules, it is possible to represent the underlying vowel system of Kasem as simply /a, i, u/, with the added vowels e and o derived from underlying sequences ai and au, respectively.<sup>2</sup> This representation, in turn, permits us to approach the most problematic looking set of forms:

- 3.7 a. koga 'back'; pl. kwe  
 b. čoŋa 'path'; pl. čwe

In 3.8 below, the first and second columns consist of the underlying representations of the singular and plural forms, respectively. The third column represents the effect of the velar elision rule in the plural forms; these, in turn, must be turned into the forms in the fourth column in order for the rules of glide formation (3.4) and guŋa (3.5) to produce the correct phonetic output:

- 3.8 a. /kaug+a/      /kaug+i/      /kau+i/      /kua+i/  
 b. /čauŋ+a/      /čauŋ+i/      /čau+i/      /čua+i/

The rule required to convert the third column to the fourth column is obviously a metathesis; of three consecutive vowels, the first two are interchanged:

- 3.9      [+ syll]    [+ syll]    [+ syll]       $\longrightarrow$       2      1      3  
                  1                   2                   3

This metathesis rule is useful in accounting for a number of different forms, some of which can be found in SPE.

Consider now the forms pia 'sheep'; pl. pe, discussed in SPE pp. 360-1. Since the plural form ends in [e] = /a+i/ and the singular in i+a, Chomsky



and Halle want to derive this pair from the stem /pia/ plus /+a/ and /+i/. We see that in /pia+i/, rule 3.9 will metathesize the stem vowels, giving /pai+i/; truncation will then reduce this to /pai/, from which 3.5 will give the correct [pe]. But what of /pia+a/? We would expect metathesis to apply to this form, also producing /pai+a/; from this form, we might get either [pea] (by 3.5) or [paya] (by 3.4), but not [pia]. If metathesis did not apply, /pia+a/ would become /pia/ by truncation, perfectly regularly. We could obtain this result if the rule of truncation could precede the metathesis rule. We saw in /pia+i/  $\rightarrow$  /pai+i/  $\rightarrow$  /pai/  $\rightarrow$  [pe], however, that truncation must apply after metathesis.

Chomsky and Halle resolve this difficulty by the entirely ad hoc device of restricting the metathesis rule to application only in the event the second and third vowels are not identical, with both being a. While it is, of course, logically possible that this is the correct solution, it seems something less than intuitively clear, especially when one realizes that the sequences to which the metathesis rule is prevented from applying are precisely those which would be eliminated by vowel truncation. The limitation of the restriction to sequences of the vowel a is in no way necessary; there is no reason on the basis of the available data not to prevent metathesis in all sequences containing adjacent identical vowels, regardless of their quality. Cases involving vowels other than a simply do not arise in the corpus available.<sup>3</sup>

The natural step of expanding the restriction on metathesis to include all sequences of identical vowels, however, is clearly a notational variant of allowing the truncation rule to apply before metathesis. The statement of this restriction as an ad hoc fact about the rule of metathesis, which obscures the relation of the restriction to another rule in the grammar,

should be avoided by permitting the truncation rule to apply generally.

Let us consider what unmarked orderings can be defined between this pair of rules (truncation and metathesis) whose ordering appears to present a paradox. In the case of /pia+i/ (or /pai+a/; cf. fn. 3), where metathesis precedes truncation, this ordering is a feeding order, and hence unmarked. In /pia+a/ (or /pai+i/), however, where truncation precedes metathesis, neither order is unmarked, or, rather, both possible orders are marked. This is because in forms of this shape, the application of metathesis will remove some possible forms that might otherwise undergo truncation, and hence the order metathesis - truncation is a bleeding order. If truncation applies first, on the other hand, it will remove some forms that might otherwise undergo metathesis, and the order truncation - metathesis is also a bleeding order. Since a decision about the relative order of these two rules is not predictable on general grounds, then, it is clear that the grammar will have to include a statement of the ordering 'truncation precedes metathesis', at least for forms of the shape  $CV_1V_2V_2$ . If the grammar does contain such a statement, however, how are we to allow derivations such as /pia+i/  $\rightarrow$  /pai+i/  $\rightarrow$  /pai/  $\rightarrow$  [pe], where the relative ordering of truncation and metathesis is predictable on universal grounds (since it is a feeding order and hence unmarked)?

A solution to this problem could be found if we were to distinguish, among the ordering statements that appear explicitly in the grammar, between some that apply absolutely and some that apply only contingently. That is, some restrictions have the property of requiring the rules they relate to apply in marked order, even though an unmarked order is available. Such statements are absolute ordering conditions. It seems reasonable that there should also be statements in the grammar of contingent restrictions, which

have the effect of determining the order of two rules only in case there is no other basis for a choice.

We have seen in this example that the question of whether a given order of rules is marked or unmarked is not always decidable from an examination of the rules alone; this notion is determinate, in general, only with respect to a specific form to which the rules are to apply. It was noted above that the structural description of a rule is really a specification of a class of forms to which the rule applies; it should come as no particular surprise that situations arise in language, where the members of a set of forms  $F_1, F_2, F_3, \dots$ , such that all are members of the classes associated with the operations  $R_1$  and  $R_2$  but for some, say  $F_1$ , the order of operations  $R_1 - R_2$  is unmarked while for others, say  $F_2$ , the opposite order is unmarked, while for others, say  $F_3$ , no unmarked order exists. Other examples of this situation were seen earlier (e.g., the u-umlaut rule in Icelandic); it will be seen that the existence of such situations makes possible a distinction between absolute and contingent ordering restrictions of the type suggested here. The intuitive basis of the distinction is that absolute restrictions are required to enforce marked orders where unmarked orders exist, while contingent orders exist only to specify an ordering where all have the same markedness status (either both marked or both unmarked or neither marked or unmarked). By founding the distinction on the notion of markedness, it is possible to predict the non-occurrence of certain kinds of situations: for example, although the notion of contingent order does associate an ordering restriction with a class of forms, this relation is defined on universal grounds, and not ad hoc for a particular language; thus, one would not expect to find a situation in which one contingent ordering restriction applied to one

(unsystematic, language particular) class of forms, while the opposite contingent order applied to another. By defining contingent orders in terms of markedness alone, and not directly in terms of language particular classes of forms, such cases are excluded, though of course such a theory could be imagined. The claim that no cases exist for which it would be required, implicit in the definition of contingent ordering, is an empirical one. We will see later that there do exist cases in which a language particular class does have to be specified as the domain of an ordering, but such cases will be limited to ones in which a language has a class of forms that are exceptional with respect to an absolute restriction; that is, while the rules in question generally have to apply in a marked order, there are a few forms to which they apply in an unmarked order.

The Kasem ordering problem, then, is resolved by stating a contingent restriction 'truncation precedes metathesis.' It may well be that further investigation will reveal general properties of such contingent restrictions, as opposed to absolute restrictions, that will make it possible to avoid explicitly characterizing all ordering relations as one or the other in the grammars of particular languages. Regardless of the persuasiveness of the above discussion of absolute versus contingent orderings, however the support this example furnishes for the elimination of the linearizability requirements (in this case, for the elimination of the requirement of asymmetry) is unimpaired; it is clear that any adequate solution to the Kasem problem will involve the rules of truncation and metathesis, applying in different orders in different forms.

V.4 Grassmann's law in Sanskrit. One of the best known rules in Sanskrit historical phonology is known as Grassmann's law, according to which the

first of two aspirate stops in a root is de-aspirated. A version of this rule applies in Greek, as well, which has posed numerous problems for Indo-Europeanists, but this issue will not be discussed here. That Grassmann's law is part of the synchronic grammar of Skt, as well as a historical event, is seen in part from reduplicated perfects such as those in 4.1:

4.1	<u>root</u>		<u>present</u>	<u>perfect</u>
a.	phal	'burst'	phalati	paphala
b.	ḍhauk	'approach'	ḍhaukati	ḍuḍhauka
c.	khād	'chew'	khādati	cakhāda

In these and many other forms, an aspirate is reduplicated as a non-aspirate. In addition, there is a class of roots with an apparently 'mobile' aspirate, which appears in the final stop of the root in some forms of the word, and on the initial stop of the root in others:

4.2	<u>root</u>		<u>present</u>	<u>future</u>
a.	budh	'understand'	bodhati	bhotsyati
b.	duh <sup>4</sup>	'milk'	duhati	dhoksyati

In forms like these, when the root occurs in an environment where the final stop would lose its aspiration, this aspiration appears in the initial. Such roots could be accounted for by positing an abstract form with two aspirates, and having it undergo Grassmann's law if the root final aspiration is not lost by some other process. That the rule must be formulated so as to require the second aspirate in its structural description to be part of the root is shown by forms like 4.3, where an aspirate in an ending does not cause the root to lose its aspiration:

4.3	a. bhr̥	'bear'; 2 pl. act. pres.	bibhr̥tha
	b. dha	'put'; 2 du. mid. pres.	dadhāthe

We suggest, then, that the synchronic form of Grassmann's Law (= GL) be formulated in some such way as this:

$$4.4 \quad [+ \text{ cons}] \rightarrow [- \text{ asp}] / \text{ \_\_\_\_\_\_ } [+ \text{ seg}]_0 \begin{bmatrix} + \text{ cons} \\ + \text{ asp} \\ + \text{ ROOT} \end{bmatrix}$$

In addition to this rule, Sanskrit contains a rule which, among other things, de-aspirates a stop before an obstruent or # boundary, as in:

4.5	<u>root</u>	<u>nom.sg.</u>	<u>acc.sg.</u>	<u>instr.pl.</u>
a.	path 'road'	pat	patham	padbhis
b.	-vr̥dh 'increasing'	vr̥t	vr̥dham	vr̥dbhis
c.	bhudh 'awakening'	bhut	budham	bhudbhis

This process may be collapsed with the rule which devoices finals and performs other sandhi operations, but for our purposes it is sufficient to formulate a de-aspiration rule (DaC):

$$4.6 \quad [+ \text{ cons}] \rightarrow [- \text{ asp}] / \text{ \_\_\_\_\_\_ } \left\{ \begin{array}{l} [+ \text{ obst}] \\ \# \end{array} \right\}$$

As we see from forms like 4.5c, where the root is of the diaspirate class discussed above, DaC must precede GL in the derivation of such forms.

The situation is more complicated, however; Sanskrit has another rule, referred to in its historical form as Bartholomae's law, according to which a cluster consisting of a voiced aspirate stop followed by a voiceless stop becomes voiced unaspirate followed by voiced aspirate. This rule applies when the root ends in a voiced aspirate and is followed by one of the endings ta (past participle); tas (3 du. act.); tum (inf.); tvā (gerund); tha (2 pl. act.); and thas, among others:

- 4.7 a. ru(n)dh+tas → runddhas  
 b. ru(n)dh+thas → runddhas  
 c. labh+ta → labdha

If Bartholomae's law did not apply in these forms, we would expect the final stop of the root to be deaspirated by DaC, and the voicing assimilation rule to devoice the final, yielding clusters like -tt-, -tth-, and -pt-, instead of the observed -ddh- and -bdh-. Of the effects of the rule,

the deaspiration of the root-final stop will be taken care of by DaC, and the synchronic Bartholomae's law need only deal with the voicing and aspiration of the suffix stop:

4.8 (BL)

$$[- \text{cont}] \rightarrow \begin{bmatrix} + \text{voi} \\ + \text{asp} \end{bmatrix} / \begin{bmatrix} + \text{voi} \\ + \text{asp} \end{bmatrix} \begin{bmatrix} \text{---} \\ - \text{voi} \end{bmatrix}$$

When we consider diaspirate roots, in forms to which BL should apply we get the following:

- 4.9 a. /bhudh+ta/ → buddha 'understanding'  
 b. /bhandh+thas/ → banddhas 'bind'

Since the initial is deaspirated, GL must apply; hence GL must precede DaC in these forms, as must BL. This seems to contradict, however, the observation that in forms like 4.2 DaC precedes GL.

The facts cited here are all well known, of course, but problems of order in synchronic grammars were not so troublesome before linguistic theory began to concern itself with their formalization. To my knowledge, no 19th century philologist was seriously concerned with the problem of ordering just noted. Whitney (1889), however, takes a different line on the problem of the diaspirate roots, a line which enables him to give an account of the problem without resulting in an ordering paradox. He assumes that the aspiration which shows up on the initial of, e.g., bhotsyati is the same segment as that which shows up in the root final of bodhati. Thus, there are no diaspirate roots, and no Grassmann's law to account for them; instead, certain roots (= the same class as those treated above as diaspirate) are exceptional in that they are subject to a minor rule<sup>5</sup> of aspiration-metathesis (AM). This rule moves the aspiration (which must, therefore, be a separate segment) from root final to root initial when it cannot remain in final position (because of the presence of a following obstruent or #,

etc.). This analysis will also incorporate another rule moving the aspiration corresponding to Bartholomae's law, which permutes aspiration from a voiced root-final around a following voiceless stop (and in the process, voices the latter). Let us call this second rule aspiration permutation (AP). If AP precedes AM, there will no longer be any aspiration present for AM to displace to the initial (if the root is one of those subject to AM). Thus, Whitney would have derivations such as:

4.10	/budh+am/	/budh+sya+ti/	/budh+ti/
	[+AM]	[+AM]	[+AM]
AP:	/budh+am/	/budh+sya+ti/	/bud+dhi/
AM:	/budh+am/	/bhud+sya+ti/	/bud+dhi/

That this is the analysis Whitney intended is shown by his statement about forms like buddhi that "In this combination, as the sonant aspiration is not lost but transferred, the restoration of the initial aspiration (sec. 155)<sup>6</sup> does not take place."<sup>7</sup>

Despite the fact that Whitney's analysis gives the correct forms, and no single ordering of the rules GL, BL, and DaC above does as well, it must be rejected as part of the grammar of Sanskrit for several reasons: 1) there is no other reason to treat the aspiration in the forms just discussed as a separate segment, and in fact doing so will complicate other parts of the grammar. Zwicky (1965) has argued (on exceedingly weak grounds, as will be discussed below) for the analysis of the voiceless stops as clusters, but no such argument is available for the voiced ones. 2) Linguistic theory must allow metathesis rules in linguistic descriptions, as was shown in the description of Kasem in section V.3, but it must constrain these highly. The rule AM seems, to this writer, highly unnatural. In addition, the rule AP must specify that the intervening stop around which the aspiration is



permuted is voiced as a result. This seems implausible. 3) Rules such as DaC must be in the grammar in any event, independently of the 'diaspirate roots'. The rule of AM will have to duplicate the statement of all such rules which might deaspirate the final as part of the statement of the conditions under which AM applies. This results in a tremendous loss of generality. 4) This solution treats the facts that aspirates are reduplicated as unaspirated and that diaspirate roots show only one aspirate in any given form as totally unrelated. Linguists since Grassmann have agreed that the impossibility of sequences of aspirates in both places is a significant generalization, which should be captured in the grammar by dealing with both as reflexes of the operation of the same rule.

The impossibility of ordering the three rules of GL, BL and DaC linearly has recently been noted by Zwicky (1965) and Kiparsky (1965). Each of these writers has attempted to circumvent the problem by reformulating the traditional rules so as to permit them to be arranged in a linear order. Though distinct, these two reformulations are closely related, and evidence will now be presented that both fail adequately to account for the facts.

Zwicky's solution rests on reformulating Grassmann's law so that it is restricted to those forms in which the second aspirate is followed by a sonorant, t, or th. Due to the restrictions which exist on the set of possible Skt endings, the only other segments which could possibly follow the root are s, dh, and bh. If the scope of Grassmann's law is restricted as Zwicky suggests, then this new rule (GL') can be ordered before DaC, since the results of DaC are never relevant to determining the applicability of GL' to any form. The resultant derivations would be:

4.11	/bhudh+am/	/bhudh+sya+ti/	/bhudh+ti/
GL':	/budh+am/	/bhudh+sya+ti/	/budh+ti/
BL:	/budh+am/	/bhudh+sya+ti/	/budh+dhi/
DaC:	/budh+am/	/bhud+sya+ti/	/bud+dhi/

Since the class consisting of the sonorants and voiceless stops, as opposed to other obstruents, is not, to say the least, the world's most natural one, Zwicky's analysis is highly suspect from the start. In order to remedy this to some extent, Zwicky proposes a re-analysis of other aspects of Sanskrit phonology, the result of which will be a system in which his rule (GL') can be formulated more naturally. Central to this effort, he suggests that the voiceless aspirates are to be analyzed as underlying clusters of voiceless stop plus h, while the voiced aspirates are simply [+ tense] voiced stops. The benefit of this to Zwicky is that now both t and th are [- tense], while dh and bh are both [+ tense]. Sonorants, of course, are all [- tense], and while there is no particular reason to mark s as [+ tense], there is no particular reason not to. Thus, all of the segments before which the rule GL' does not apply can be marked [+ tense], resulting in a more natural looking rule. The full set of Zwicky's rules, in the order they are to be applied, is given here in 4.12, together with the rules which form the voiceless aspirates (after cluster representation is no longer needed) and specify the tense aspirates as voiced:

- 4.12 a. (GL') [- cont]  $\rightarrow$  [- tense] / \_\_\_\_ [+ cont] <sub>1</sub>  $\begin{bmatrix} - \text{cont} \\ + \text{tense} \\ + \text{ROOT} \end{bmatrix}$  [-tense]
- b. (BL') [+ obst]  $\rightarrow$  [+ tense] / [+ tense] + \_\_\_\_
- c. [+ tense]  $\rightarrow$  [+ voice] /  $\begin{bmatrix} - \\ - \text{cont} \end{bmatrix}$
- d.  $\begin{matrix} [- \text{cont}] & h \\ 1 & 2 \end{matrix} \rightarrow \begin{bmatrix} 1 \\ + \text{tense} \end{bmatrix} \quad \emptyset$
- e. (DaC; same as 4.6 above)

While these rules, like Whitney's, work, they do not stand out as blindingly natural. In particular, the claim of GL' that the tenseness of the segment following two aspirates determines the aspiration of the first of them seems wrong. Also, since the voiceless aspirates are no longer tense at the time GL' applies, it cannot affect them, and the fact that such stops are reduplicated as unaspirated must then be due to some other fact. It is true that the reduplication rule takes only the first element of a cluster (except that it takes the stop of an s+stop sequence), and hence if the voiceless aspirates are represented as cluster, the correct output will emerge. It will still be true that the generalization 'aspirates are preceded by unaspirates' will be divided between at least two rules, rather than being stated once as GL.

The crucial assumption behind Zwicky's analysis is that the voiceless aspirates can be represented as clusters. Aside from the fact that this would make the rules in 4.12 work (which is presumably counterbalanced by the otherwise unnecessary rule 4.12d), the only argument given for this position is a very tenuous one, based on the appearance of a linking vowel before certain consonantal endings. It is very difficult to make any consistent generalizations about the appearance of this vowel; however, there are three classes of roots which generally take it (among others): 1) roots ending in a cluster of consonants; 2) roots ending in a retroflex (except the ʃ from ʃ after r, u, k, or i); and 3) roots ending in a voiceless aspirate. Zwicky has given (elsewhere in the same work) a rather shaky argument that root-final retroflexes should come from clusters of ɭ plus a dental. If this is accepted, then the re-analysis of voiceless aspirates as clusters of stop plus h would permit the three classes listed above to be reduced to the one class 'roots ending in a consonant cluster.' Even

if all of the steps of this argument are accepted, however, we are a long way from explaining all of the occurrences of linking vowels, and it can hardly be counted as strong evidence.

More important than the above considerations, however, is the fact that there are some forms in Skt which simply cannot be accounted for at all under Zwicky's analysis. In particular, there is a stem dhadh (from the root dhā) which, among its other anomalous properties, is traditionally described as being exceptional in that it does not undergo Bartholomae's law. Forms of this stem are:

- 4.13 Stem: dadha-; 3 sg.act.pres.: dadhāti; 2 pl.act.pres.: dhattha; 3 du.act.pres.: dhattas; 2 pl.imperf.mid.: ~~adhaddhvam~~ (from /adhadh+dhvam/)

We see in this form that the aspiration of the second dh disappears before all obstruents, in which case the first appears as surface dh. Before the endings which begin with t or th, the root-final stop is deaspirated and devoiced, instead of being subject to Bartholomae's law. In order to account for a form like dhattha on the basis of Zwicky's analysis, we must say that the root dhadh is exceptional with respect to both GL' and BL'. If it undergoes GL', it will appear incorrectly as dattha, while if it undergoes BL', it will appear with medial -ddh- instead of -tth-. This double exceptionality is already a liability; it is impossible for a form to be exceptional simply with respect to Bartholomae's law; two exception features are required to achieve this effect. Especially in early texts like the Rgveda, many forms are encountered which do not show the effect of BL where the conditions for it are met, but all of these behave like dhadh in being exceptional (on Zwicky's analysis) to both GL' and BL'. No forms are encountered which are exceptional with respect to only one of

these rules, which strongly suggests that there is only one rule involved. More importantly, however, if roots like dhadh are marked as exceptional with respect to GL', as must be the case if we are to get dhattha, then it will be impossible to get dadhāti, a form to which GL' must apply. There is thus no way at all to account for the behavior of dhadh, whether on principled or on other grounds. Since the complex of facts related to dhadh is not restricted to this one root (which might, after all, be treated as purely suppletive), but is rather a fairly common source of irregular doublets in archaic texts, the inability of Zwicky's analysis to encompass it at all is a serious defect.

It will be noticed that the spirit of Zwicky's solution is similar to Whitney's, in that he prevents GL from applying in precisely the cases where DaC but not BL will apply by building part of the apparatus of these rules into the formulation of GL'. Kiparsky's (1965) analysis is very similar, but he takes a slightly different line. Kiparsky proposes to leave GL as formulated above (4.4), but to restrict DaC so that it does not apply before t or th. He does not discuss the formal statement of the resulting rule, and it is not immediately obvious that a plausible formulation is possible. It is again necessary to distinguish the class of sonorants and voiceless stops from the class containing s, dh, and bh. Whatever problems beset Zwicky's analysis on this score will probably be Kiparsky's as well. In any case, Kiparsky is further called upon to invoke the principle of cyclic application, a principle which has (to date) no support in the domain of segmental phonology. Kiparsky's rules, and the typical derivations they give rise to, are as follows:

$$4.14 \text{ a. (DaC'')} \quad [+ \text{ cons}] \rightarrow [- \text{ tense}] / \text{ \_\_\_\_\_\_ } \{s, dh, bh, \# \}$$

(4.14 continued)

b. (GL) [+ cons]  $\rightarrow$  [- tense] / \_\_\_\_ [+ cont] <sub>o</sub>  $\left[ \begin{array}{l} - \text{syll} \\ + \text{tense} \\ + \text{ROOT} \end{array} \right]$

c. (BL) [+obst]  $\rightarrow$   $\left[ \begin{array}{l} + \text{tense} \\ + \text{voice} \end{array} \right]$  /  $\left[ \begin{array}{l} + \text{obst} \\ + \text{tense} \\ + \text{voice} \end{array} \right]$  \_\_\_\_

		/dhogh+ti/	/dhugh+dhve/	/adhogh+am/
first cycle	a)	"	/dhug+dhve/	"
	b)	/dogh+ti/	"	/adogh+am/
	c)	/dogh+dhi/	"	"
<hr/>				
second cycle	a)	/dog+dhi/	"	"
	b)	"	"	"
	c)	"	"	"
<hr/>				
Output:		[dogdhi]	[dhugdhve]	[adoham]

Observe that restricting DaC so that it does not apply before t or th could never have the effect of preserving a segment's aspiration; in every case, the t or th will become d or dh when BL comes around, and DaC will then apply on the next cycle. It is not clear, therefore, how this restriction on the rule could ever be incorporated into it in the process of language learning. Notice also that the domain of the cycle is the same on both passes; the only effect of the cycle is to obtain a different order of application of 4.14a and 4.14b in those forms to which 4.14c does not apply. Kiparsky's analysis must be regarded as labored and overly complex. But notice that, in addition, the stem dhadh cannot be accounted for on this analysis any better than it could on Zwicky's. In order to generate dhattha, it is necessary not only that 4.14c not apply (exceptionally), when it should, but also that 4.14a apply (exceptionally) where it should

not. This latter kind of exception runs the risk of complete incoherence. But even if dhadh is marked in some (hitherto undefined) way so that 4.14a will apply to it even if its analysis is not satisfied, it will then become impossible to generate dadhāti, since the (incorrect) application of 4.14a here will give the incorrect \*dhadāti. Thus, Kiparsky's analysis is subject to the objections raised against Zwicky's, and is, like it, to be rejected.

We see that all of the ways that have been proposed (by Whitney, Zwicky, and Kiparsky) to resolve the problem of the interrelationship of GL, BL and DaC fail in one way or another. The problem in the traditional analysis, as has been pointed out, is that the rules GL and DaC apply in different orders, depending on whether the form is also subject to BL. In a theory of ordering admitting only linearizable ordering relationships, such a situation is of course impossible, and requires that the rules be reformulated. Since such reformulation, however, seems out of the question if we are to preserve the required generalizations about the structure of Sanskrit, the place to look for an alternative resolution is in the constraints imposed on the ordering relation.

In seeking a solution under the ordering theory being defended here, let us begin by observing that the forms of GL and BL are such that their interaction does not define a marked or an unmarked order between them. Let us then propose the restriction 'GL precedes BL' as part of the grammar of Sanskrit. Now, given a form like /bhudh+ta/, we see that the order 'BL precedes DaC' is unmarked, for the opposite order is, in fact, a bleeding order which would eliminate all possible applications of BL from the language. Given that BL is a rule of Sanskrit, then, it must apply in this order with respect to DaC. Now in a form like /bhudh+ta/, the rules will apply in the order GL - BL - DaC, resulting in the correct output [buddha]. This is the

only order which is consistent with the requirement 'GL precedes BL', and also with the unmarked order 'BL precedes DaC.'

Now consider a form like /bhodh+sya+ti/. Rule BL cannot apply to this form, so the restriction 'GL precedes BL' cannot be relevant here. The two rules that could apply are GL and DaC. It appears that the order 'GL precedes DaC' is an unmarked one, since the opposite order removes representations to which GL could apply. Recall, however, the original basis of the notion 'unmarked order': this was 'the order that maximizes the applicability of the rules.' Though it appears that 'GL - DaC' is such an order, consider the consequences of applying these rules in this way: the initial aspirate of every occurrence of a diaspirate root would be de-aspirated; as a result, the class of diaspirate roots would cease to exist. That is, these roots would all be reanalyzed as ordinary roots with unaspirated initials and aspirated finals. Once this occurred, there would no longer be any reason for a rule like GL to be in the grammar, since its only remaining cases would be treated as a special fact about reduplication. Thus, while it appears that the order GL - DaC is one leading to wider applicability for GL, in fact such an order would lead to the complete elimination of GL from the grammar. Given the rest of the grammar (and lexicon) of the language, the fact that GL is a rule of the grammar means that it would be (logically) impossible for it to always apply before DaC, given also a reasonable set of assumptions about restructuring. We are led to conclude, therefore, that the truly unmarked order of GL and DaC is DaC - GL, and that the only forms in which the other order will be observed are those in which some explicit requirement leads to this. This is exactly what is observed: in the forms like /bhodh+sya+ti/, where only GL and DaC applied, they applied in the unmarked order DaC - GL. In a form like /bhudh+ta/, however, the explicit ordering



relationship 'GL precedes BL' enforces another relative order on GL and DaC. Since this order does not contradict any explicit requirement of the grammar, it is perfectly permissible. It is only the fact that the normal case, to which many forms (like /bhodh+sya+ti/) are subject, has the order which preserves the existence of the rule GL that permits other cases to exist with the other order of these two rules.

Notice that while the ordering statement 'GL precedes BL' seems arbitrary, it is in fact dictated by the facts of the language. That is, assume we chose instead 'BL precedes GL'. Now BL - DaC has been seen to be unmarked; hence both DaC and GL follow BL. The unmarked order of DaC and GL has just been seen to be DaC - GL; hence the only possible order of the three rules would be BL - DaC - GL. The result of applying the rules in this order, however, would be the incorrect \*bhuddha. The facts, therefore, require that the order of BL and GL be GL - BL in order to enforce the correct order on GL and DaC in forms to which BL applies.

Notice that the stem dhadh is no longer a problem. It is simply marked [-BL]; the non-applicability of BL will then have the consequence that a form like dhattha will be treated exactly like any other form to which GL and DaC, but not BL, apply; that is, they will apply in the (unmarked) order DaC - GL, as required.

The solution which arises from the framework of assumptions about ordering relations adopted here, then, preserves the intuitively correct and generally agreed upon form of the important generalizations about the structure of Sanskrit, without giving rise to a paradox. It also avoids the defects which the solutions based on reformulations of the rules were found to have. As such, it constitutes significant evidence for the theory of local ordering, which makes it possible, and for the relaxation of the requirement of linearizability.

V.5 Turkish harmony. One of the best known phonological rules in any language is the group of processes collectively called vowel (and consonant) harmony in Turkish. Frequently cited as the very type of a phonological rule, the harmony processes in Turkish have a very general applicability, and due to the agglutinative character of the language, are evidenced in virtually every word. The descriptions in the traditional grammars of Turkish (e.g., Swift, 1963, and Kreider, 1954) have tended to be rather sketchy, generally consisting of little more than a list of the possible sequences of vowels found internal to words and morphemes. Lees (1961; 1967) has discussed the problem within the framework of generative phonology; a recent discussion by Foster (1968) has organized all of the facts that appear relevant, and forms the basis of this section.

The facts, in outline, are as follows. Turkish words, which consist of a stem and a number of suffixes, contain in general only vowels with the same value for the feature [ $\pm$  back]; and furthermore, a high vowel in the second or a later syllable has the same value for the feature [ $\pm$  round] as the vowel in the preceding syllable. Non-high vowels after the first syllable are always [- round], a common characteristic of Altaic languages. In addition, the velars /k, g/ have two shapes: [+ back] in words with [+ back] vowels, and [- back] in words with [- back] vowels. The lateral liquid /l/ also has two allophones: a [+ back], 'dark l' in words with [+ back] vowels, and a [- back], 'light l' (which may or may not be [+ high], or palatalized as well) in words with [- back] vowels.<sup>8</sup> As a result, most suffixes in the language appear in at least two phonetic shapes, depending on the values of the harmonizing features in the stem.

One of the first issues which must be resolved for the description of this process is whether or not 'consonant harmony' (i.e., the alternations

in /k, g, l/) is to be accounted for as part of the same rule as 'vowel harmony.' Since this harmony involves one of the same features as does vowel harmony ([ $\pm$  back]), there is a strong temptation to seek a single formulation of the class of segments affected by 'backness harmony.' Foster (1968), however, argues that the two should not be treated by the same rule. His first point in favor of separating them is an argument from universal phonetics. The assimilation of backness in velars to that of surrounding vowels is a very general, if not universal phenomenon; the assimilation of the backness of a vowel to that of a neighboring vowel is restricted to a much smaller class of languages. Since the one process is frequently found without the other, they should not be identified where they are found together, as in Turkish.

The plausibility of this argument, taken by itself, is quite low. For one thing, it should be noted that the assimilation of l in the same way as the velars is not as widely distributed. More common is the velarization of /l/ before a consonant or # boundary, with 'light l' showing up intervocally and perhaps initially. Thus, since it affects /l/ as well as /k, g/, the consonant harmony rule will have to be a language particular rule, like the vowel harmony rule. But further, as will be pointed out in chapter VI, the language-specific realizations of universal phonetic processes often have very particular conditions which are inseparable from the general ones. The argument from universality of the velar assimilation, then, has little force.

Other facts adduced by Foster, however, are more indicative of the status of consonant harmony. He points out that there are words like 5.1a in which the velars, exceptionally, do not agree in backness with the neighboring vowels, but which behave normally with respect to vowel harmony, as is shown



words, it is generally true that the last vowel of the stem determines the backness of the suffix vowels:

- 5.4 a. ziyaret 'visit'; ziyaretinizden 'because of your (pl.) visit'  
 b. kitab 'book'; kitablar 'books'  
 c. eşya 'collection of items'; eşyalar 'collections of items'  
 d. şeftali 'peach'; şeftalilerin 'peaches (gen. pl.)'

Since the backness harmony rule does not apply within these non-harmonic stems, but does apply between their last vowel and the vowel of a suffix, we conclude that the harmony rule requires the presence of a morpheme boundary in its analysis.<sup>11</sup> Furthermore, there are some suffixes that contain vowels not subject to the operation of the rule; an example is the suffix /Iyor/, whose 'o' is not subject to harmony processes, but is invariant. When this suffix appears with a front-vowel stem, such as /gel-/ 'come', its initial vowel is marked [- back], as expected; the 'o' is unaffected; and following vowels, such as the vowel of the suffix /-Im/ 'first person' take the backness of the o (i.e., become [+ back]):

5.5 /gel+Iyor+Im/ → [g<sub>1</sub>el<sub>1</sub>iyorum] 'I am coming'

Since the backness of a vowel affects only those vowels that follow it, we may assume the assimilation is progressive and not regressive. As we saw above, consonant harmony operates in both directions; this typological difference between the rules is further evidence of their separability. We might now formulate the rule either as an infinite schema or as an iterative process (as described above in section IV.3). It seems better to write the rule as iterative, however. Recall that certain vowels are non-harmonic, and that a given 'harmonic domain' extends just up to such a vowel, and the string following it is another 'harmonic domain.' If this is formulated as an infinite schema, it will be necessary to assign these vowels an arbitrary

diacritic feature, say [- HARM], and make the rule sensitive to this feature:

$$5.6 \quad [+ \text{syll}] \longrightarrow [\alpha \text{ back}] / \left[ \begin{array}{c} + \text{syll} \\ \alpha \text{ back} \end{array} \right] C_o + (C_o \left[ \begin{array}{c} + \text{syll} \\ + \text{HARM} \end{array} \right])_o C_o \text{ ---}$$

Notice that the feature [ $\pm$  HARM] has no direct interpretation; every vowel must be marked for it, and the fact that its function is purely and simply to prevent the harmony rule from applying is unexpressed. There is no reason why other rules of the grammar could not change it, or at least be sensitive to it. Current research in generative phonology seems to indicate that such a powerful device for dealing with exceptions is not needed otherwise; another solution to this problem would certainly be preferable. If we write this rule as an iterative schema, however, such a device is not necessary; it is never necessary, in this case, to delimit the range of a given harmony, but simply to indicate for a given vowel that it is exceptional with respect to a given rule. The use of such specific exception features, inherently restricted to determining the applicability of a single rule, are well established devices in phonology, as opposed to the greater power of purely arbitrary exception features. We can thus write the backness harmony rule as follows:

5.7 (iterative)

$$[+ \text{syll}] \longrightarrow [\alpha \text{ back}] / \left[ \begin{array}{c} + \text{syll} \\ \alpha \text{ back} \end{array} \right] C_o + C_o \text{ ---}$$

We can ask, concerning the formulation of rule 5.7, whether it is correct to express it as an  $\alpha$ -rule, or whether it might not be the case that all suffixes have a single value for the feature [ $\pm$  back], and that the harmony rule simply produces the other value when this is what is required. In fact, there is some reason to believe that this latter suggestion is correct; there are a large number of words in Turkish that

should, apparently, take back vowel suffixes, but instead take front vowel suffixes:<sup>12</sup>

- 5.8 a. usul 'system'; usulüm 'my system'  
 b. alkol 'alcohol'; alkollu 'his alcohol'  
 c. helal<sub>1</sub> 'legitimate property'; helal<sub>1</sub>in 'your leg. property'  
 d. kalb 'heart'; kalbim 'my heart'  
 e. k<sub>1</sub>ab 'ankle'; k<sub>1</sub>abi 'her ankle'  
 f. harf 'letter of the alphabet'; harfler 'letters'  
 g. kabahat 'fault'; kabahatti 'his fault'  
 h. saat 'watch, time'; saatim 'my watch'  
 i. imsak 'fasting'; imsakkinden 'on account of his fasting'

A large number of these words end in (exceptional) l<sub>1</sub>; traditional accounts of the phenomenon have generally attempted to establish this as the feature conditioning the suffix vowels. As the examples above show, however, the phenomenon is by no means confined to words in final -l<sub>1</sub>. Foster has noted this, and suggested that in fact any consonant (at least in final position) can be palatalized, with this feature later eliminated from phonetic representations for all consonants except [k<sub>1</sub>, g<sub>1</sub>, l<sub>1</sub>]. The backness harmony rule, then, will be sensitive to the backness of either a final vowel or a final consonant of the stem. This seems a serious case of throwing out the baby with the bathwater; the feature posited for consonants has no supporting evidence whatsoever, and there are no other grounds, universal or particular, that would lend it plausibility.

Suppose, however, that the suffixes have vowels that are represented as [- back] in underlying forms, and the harmony rule simply makes certain vowels [+ back]. Then the stems in 5.8 can simply be marked as exceptions to the harmony rule, in that following suffixes are not affected. In that case, there will be no reason either to alter the representations of their final consonants or to reformulate the rule so that it is sensitive to either

a vowel or a consonant. On this analysis, the backness harmony rule will be as follows:

5.9 (iterative)

$$[+ \text{syll}] \rightarrow [+ \text{back}] / \begin{bmatrix} + \text{syll} \\ + \text{back} \end{bmatrix} C_o + C_o \text{ —}$$

The remaining harmonic process to be accounted for is that involving the feature [ $\pm$  round]. It will be recalled that only high vowels participate in this harmony, since non-high vowels after the first syllable (in regular words) are always non-round. We can thus write a single rule to unround vowels after the first syllable:

$$5.10 \quad [+ \text{syll}] \rightarrow [- \text{round}] / V C_o \begin{bmatrix} \text{ — } \\ - \text{ high} \end{bmatrix}$$

Such a rule will be unnecessary, however, if we simply represent all suffixes as having the underlying shape with a [- round] vowel; in that case, a vowel which would otherwise be exceptional to 5.10, such as the o of /-Iyor/, is simply one that is [+ round] in the lexicon.<sup>13</sup> We can now formulate the rounding harmony rule as follows:

$$5.11 \quad [+ \text{syll}] \rightarrow [+ \text{round}] / \begin{bmatrix} + \text{syll} \\ + \text{round} \end{bmatrix} C_o + C_o \begin{bmatrix} \text{ — } \\ + \text{ high} \end{bmatrix}$$

This rule will also be an iterative one. It is possible to collapse it with the backness harmony rule (5.9), though the resultant saving is minimal:

5.12 (iterative)

$$[+ \text{syll}] \rightarrow \begin{bmatrix} \langle + \text{round} \rangle_c \\ \langle + \text{back} \rangle_e \end{bmatrix} / \begin{bmatrix} + \text{syll} \\ \langle + \text{round} \rangle_b \\ \langle + \text{back} \rangle_d \end{bmatrix} C_o + C_o \begin{bmatrix} \text{ — } \\ \langle + \text{high} \rangle_a \end{bmatrix}$$

Conditions: if a and b, then c; if d, then e

Rule 5.12, then, represents Turkish vowel harmony, while 5.3 above represents the (distinct) process of consonant harmony.



There are two sets of exceptions to the above rules that can be predicted systematically. The first such class is that of forms subject to so-called 'labial attraction.' This is a process whereby the second vowel of a stem may be rounded even though the first is unround, if the second is high, the first is a, and a consonant cluster including at least one labial intervenes. Such stems are the following:

- 5.13 a. *şabuk* 'quick'  
 b. *şamuka* 'sand-smelt'  
 c. *hamul* 'patient'  
 d. *maymun* 'monkey'  
 e. *samsun* 'mastiff'  
 f. *havruz* 'chamberpot'  
 g. *karmuk* 'grappling iron'

This process is confined to stem-internal vowels, and does not apply across morpheme boundaries; it is thus to be represented by a lexical structure rule such as:

$$5.14 \quad [+ \text{syll}] \longrightarrow [+ \text{round}] / C_o \underline{a} C_o \begin{bmatrix} + \text{ cons} \\ + \text{ ant} \\ - \text{ cor} \end{bmatrix} C_o \begin{bmatrix} \text{---} \\ + \text{ high} \end{bmatrix}$$

The rounding of suffixes follows properly from the rounding of the final vowels of these stems.

A more difficult set of exceptions, however, is the result of the process called by Lees (1967) 'palatal umlaut.' This phenomenon varies somewhat with the dialect under consideration; however,

"The essence of the palatal umlaut of the Istanbul dialect is the following: a short vowel is unrounded immediately before a palatal /y, ɟ, j, ç, c/ within word boundaries if morpheme-final or if not in the first syllable of the word and is, moreover, also raised there if that palatal is followed immediately by a vowel. Thus, we would have the following pronunciations:

/ye+yen/	[yiyen]
/oku+muş/	[okumış]
/üşü+yüş/	[üşiyiş]
/üşü+me+yiş/	[üşümiyiş]
/gümüş+tür/	[gümiştir]

From the last example we see that the regular vowel-harmony rules assimilate vowels to a preceding vowel after the latter has already undergone the palatal assimilation, and therefore that the rule for rounding must follow the palatal umlaut rule."

(Lees, 1967, pp. 289-90)

The segments which produce the applications of the rule are all [+ high, - back]; unfortunately, so are /k<sub>1</sub>, g<sub>1</sub>, (and perhaps) l<sub>1</sub>/. We could perhaps claim that y in Turkish is a palatal-alveolar glide, rather than a palatal; in that case, the class required would be that of [+ high, - back, + cor] segments. There is no particular phonetic motivation for this move, however; its assumption in the formulation of the rule below cannot be regarded as more than a notational device:

5.15

$$[+ \text{syll}] \rightarrow \left[ \begin{array}{c} - \text{round} \\ \langle + \text{high} \rangle_d \end{array} \right] / \begin{array}{c} \langle \# \rangle_c \\ a_a o \end{array} \text{ — } \begin{array}{c} \langle + \rangle_b \\ b_b \end{array} \left[ \begin{array}{c} + \text{high} \\ - \text{back} \\ + \text{cor} \end{array} \right] \langle [+ \text{syll}] \rangle_c$$

Conditions: if a, then b; if c, then d.

Recall Lees' statement quoted above to the effect that 5.15 must precede rounding harmony (the part of 5.12 represented by 5.11), because the unrounding of vowels by 5.15 results in unrounded suffix vowels where rounded ones would otherwise be expected. Note that this could not be incorporated simply as a condition on 5.12, for the following reasons. First, the rule 5.15 is needed in the grammar in any event, to account for the unrounding of vowels such as the second vowel of [üşiyiş] which would otherwise be round. To impose an extra condition on 5.12 would require the complete restatement

of rule 5.15 as a part of 5.12, and would result in the loss of the generalization that vowels after a palatal-umlauted vowel are [- round] because that vowel has been unrounded by 5.15. Furthermore, there are dialects, such as one discussed by Foster (1968), in which the palatal umlaut rule only raises the vowel, and does not unround it; in these dialects, the suffixes that follow palatal-umlauted vowels are still round if the basic umlauted vowel was roundable. In this dialect, then, /üşü+me+yiş/ gives not [üşümiyiş], as in the dialect described by Lees to which 5.15 applies, but rather [üşümüyüş]. If the unrounding of vowels affected by 5.15 and the unround vowels of suffixes following such vowels were separate facts, as would be claimed by stating the former as 5.15 and the latter as a condition on 5.12, we would expect them to vary independently; instead we find that the only apparent possibilities are both (as in Istanbul) or neither (as in the North Anatolian described by Foster). We must therefore assume that the [- round] quality of the suffixes in the words cited by Lees in the quote above is a consequence of the application of rounding harmony to these suffixes after the vowels affected by 5.15 have already been unrounded.

If, however, rounding harmony applies to the output of palatal umlaut, we will still get the wrong results. In this case, the vowel which was unrounded by palatal umlaut will simply be rounded again by rounding harmony, and the following suffixes will be rounded as well. Thus, the Foster dialect will be the only one generable. It seems impossible to order 5.12 and 5.15 in either order -- if they apply in the order 5.12 - 5.15, we will get [üşiyüş] instead of [üşiyiş], while if they apply in the order 5.15 - 5.12, we will get [üşüyüş].

Lees solved this problem of ordering (in Lees, 1967) by applying the rules in the order 5.15 - 5.12, and having 5.15 introduce an ad hoc diacritic

'[- rounding harmony]', which then prevented 5.12 from applying. Permitting rules to introduce diacritic features is a device whose power cannot readily be imagined; Chomsky and Halle (1968) suggested that it might be possible for rules to introduce the feature [- next rule], but this device has been called into question as excessively powerful. In any case, the feature [- next rule] would have to be assigned not to the whole morpheme, as Chomsky and Halle suggest, but only to the particular segment affected by 5.15. A solution that does not involve rules introducing diacritic features should be found, if at all possible.

One such solution that could be presented rests on the principle of extended disjunctive ordering suggested above in section IV.5. Since the class of forms in which a vowel which should be rounded by 5.12 should also be unrounded by 5.15 is a proper subset of the class of forms which should simply be rounded by 5.12, and since 5.12 and 5.15 both specify the value of the feature [ $\pm$  round], it might be possible to invoke this principle to explain the fact that 5.15 (the more specific rule) applies to the exclusion of 5.12. This solution to the ordering problem remains a possibility, but the uncertainties surrounding the extension of the principle of disjunctive ordering suggested in IV.5 make it an insecure one.

Another possibility for a solution to this ordering paradox involves crucially the formulation of rounding harmony as an iterative rule, rather than as an infinite schema. If the rounding harmony rule applies to its own output, progressively working its way across the word, then each vowel's rounding is due to a separate application of 5.12. If we allowed 5.15 to apply between two applications of 5.12, the correct results could be obtained. This follows from the fact that the unmarked order of 5.12 and 5.15 for any particular vowel must be 5.12 - 5.15, since it is only the application of

5.12 that produces a rounded vowel that can be affected by 5.15. This must therefore be a feeding order. Once 5.15 has applied to the vowel, it would appear that the unmarked state would be to permit 5.12 to re-apply, rounding the vowel once again. Observe, however, that this would nullify the effect of the [-round] portion of 5.15, resulting in the elimination of this sub-rule from the grammar. It is thus apparent that if rules 5.12 and 5.15 are in the grammar at all, the unmarked order of their application to a given vowel must be 5.12 - 5.15, with no subsequent re-application of 5.12. We have, then, derivations like these:

5.16 a.        /oku+mi<sub>ɣ</sub>+t+in+iz/

5.12: /oku+mu<sub>ɣ</sub>+t+in+iz/

5.15: /oku+m<sub>ɣ</sub>+t+in+iz/

5.12: /oku+mi<sub>ɣ</sub>+t+in+iz/

5.12: /oku+m<sub>ɣ</sub>+t+in+iz/

b.            /u<sub>ɣ</sub>u+yi<sub>ɣ</sub>/

5.15: /u<sub>ɣ</sub>i+yi<sub>ɣ</sub>/

c.            /u<sub>ɣ</sub>u+me+yi<sub>ɣ</sub>/

5.15: /u<sub>ɣ</sub>u+mi+yi<sub>ɣ</sub>/

d.            /çocuk+si+lik/

5.12: /çocuk+su+lik/

5.12: /çocuk+su+luk/

The implications of this analysis for the principle of linearizability are clear: aside from the fact that iterative application in general violates the principle of irreflexivity, the possibility of 5.15 applying between two (potential) applications of 5.12 (e.g., in 5.16a, to the vowel of /+mis+/ and to the vowel of /+in+/) violates the principle of asymmetry.

Notice that what is important is not the application of 5.12 to /+mis+/ and /+in+/ as a backing rule, but the possible application of it as a rounding rule. The formulation of 5.12 as an iterative rule, then, together with the possibility of applying another rule between two possible applications of it, yields a solution to the ordering problem considered here without an appeal to either diacritic features (as in Lees' solution) or an extension of disjunctive ordering.

V.6 Chafe's theory of 'persistent rules'. In a recent paper, Chafe (1968) has given two other examples of phonological problems whose solution appears to violate the condition of linearizability, to which he gives a different interpretation than that assumed here. The problems are repeated here for convenience.

The first set of rules is from Caddo, an American Indian language. The first rule (Chafe's C1) stresses a vowel before a 'resonant' (= nasal or glide) followed by a stressed vowel:

$$6.1 \quad [+ \text{syll}] \longrightarrow [+ \text{stress}] / \text{ \_\_\_\_\_\_ } \begin{bmatrix} - \text{syll} \\ \alpha \text{ nas} \\ \alpha \text{ cons} \end{bmatrix} \begin{bmatrix} + \text{syll} \\ + \text{stress} \end{bmatrix}$$

The second rule, which Chafe does not specify in detail, deletes certain internal unstressed vowels. When these two apply to a form like cihawáynikah 'we have run away', rule 6.1 applies first to stress the second vowel, resulting in the inapplicability of the vowel dropping rule ('C2') to this vowel. Thus, we obtain cihawáynikah. This shows that, for any given vowel, 6.1 must precede C2.

Another rule that interacts with these two converts a sequence of vowel plus h into a stressed vowel in the position before two non-syllabic segments:

$$\begin{array}{ccccccc}
 6.2 & [+ \text{ syll}] & h & [- \text{ syll}] & [- \text{ syll}] & \longrightarrow & 1 \quad \emptyset \quad 3 \quad 4 \\
 & 1 & 2 & 3 & 4 & & [+ \text{ stress}]
 \end{array}$$

Given a form like hanahya?ah 'there are many', the vowel-dropping rule (C2) applies to the third vowel, yielding hanahy?ah; this, in turn, is affected by rule 6.2 above, yielding hanáy?ah, which is then subject to 6.1, yielding the form hánáy?ah. Thus, although for any one vowel rule 6.1 must precede C2, there are forms in which C2 must precede 6.2, which in turn precedes 6.1. It will be seen that 6.1 and C2 are so related that either ordering of them (alone) is a bleeding order; if 6.1 applies first, some vowels that could be deleted will be stressed, and thereby rendered ineligible for deletion; but if deletion applies first, some vowels that could be stressed will be deleted. Thus, it is not possible to define an unmarked order between these rules. The ordering 6.1 - C2, which is necessary to generate ciháyáynikah instead of the incorrect cihwáyáynikah thus appears in the grammar as an explicit contingent ordering.

In a form like hanahya?ah, the application of C2 produces new forms for 6.2 to apply to that would not be affected if 6.2 applied first. Thus, the ordering C2 - 6.2 is a feeding order, and unmarked. Similarly, 6.2 creates stressed vowels that permit the environment of 6.1 to be satisfied, which would not be the case if 6.1 applied before 6.2. Hence the ordering 6.2 - 6.1 is feeding, and unmarked. The fact that C2 precedes 6.1 in this form, then, follows from the fact that unmarked orders are available here to define this ordering, over-riding the contingent ordering requirement 6.1 - C2. The case is formally similar to the Kasem case discussed in section V.3 above.

Chafe's other set of rules is attributed to Schane, and applies in French. The rules are as follows:

- 6.3 a.  $\emptyset \rightarrow \emptyset / C \begin{bmatrix} + \text{ cons} \\ + \text{ son} \end{bmatrix} \text{ \_\_\_\_\_\_}$   
 b.  $C \rightarrow \emptyset / \text{ \_\_\_\_\_\_ } + C$   
 c.  $[+ \text{ syll}] \rightarrow \emptyset / + \text{ \_\_\_\_\_\_ } +^{14}$

In the underlying form /ekriv+r/, both 6.3a and 6.3b could apply, but if 6.3b applies first it will make 6.3a inapplicable. The order 6.3b - 6.3a is thus a bleeding order for this form, and the opposite order (6.3a - 6.3b) is the unmarked order. The resultant form is /ekrir $\emptyset$ /. In a form such as /viv+t+r/, the v is not lost; hence, the rule 6.3c must not apply before 6.3b. This is a marked ordering, and must be stated in the grammar. After 6.3c does apply, however, 6.3a can apply (the order 6.3c - 6.3a being unmarked). Hence, for this form, the rules apply in the order 6.3b - 6.3c - 6.3a; in /ekriv+r/, they apply in the order 6.3a - 6.3b. This result can be obtained, on the theory of ordering adopted here, by the single statement of the marked order '6.3b precedes 6.3c'.

Both of these examples obviously violate the requirement of transitivity, and Chafe's purpose in presenting them was to suggest the need for a modification of the standard theory of linear ordering of rules. His proposal is as follows: the rules of a grammar are organized into several sets or 'depths'. The underlying forms are operated on first by all of the rules of the greatest 'depth' applying simultaneously; the output is then operated on simultaneously by all of the rules of the next 'depth,' and so on until the rules of depth 1 are reached, which result in the output form. All of the rules of a given depth in Chafe's examples are rules whose interaction is irrelevant to their operation; it is not clear whether Chafe means to impose this requirement on grammars, or whether he proposes to allow any arbitrary set of rules to be applied simultaneously by appearing at the same depth. Since he presents



no evidence that this kind of power is needed for grammars, we can assume that he intended to impose the requirement that the rules appearing together at any given 'depth' be such that their relative order is irrelevant to the result. If this requirement is imposed, Chafe's theory so far is nothing but a notational variant of the standard theory, with the 'vacuous orderings' eliminated. Since a linear order can always be given within each depth (by the requirement of irrelevancy), and the depths themselves are linearly ordered, the grammar is always consistent with at least one linear order of the rules.

Where Chafe's proposal departs from the requirement of linearizability is in its notion of 'persistent rule.' Chafe proposes that a certain subset of the rules at depth 1 are specially marked as 'persistend', and that a persistent rule can apply not only at depth 1, but also at various other depths. Thus, in his presentation of the Caddo example, the rule 6.1 is designated a persistent rule, and appears at depth 4 as well as at depth 1. The arrangement of the rules is as follows:

6.4    depth 4: 6.1  
          depth 3: C2 (vowel deletion)  
          depth 2: 6.2  
          depth 1: 6.1

Similarly, in the French example, the rules are arranged as follows:

6.5    depth 4: 6.3a  
          depth 3: 6.3b  
          depth 2: 6.3c  
          depth 1: 6.3a

Chafe's theory then allows a certain class of solutions which violate the requirement of linearizability. He imposes the following conditions on such violations of linearity: first, the solution must deviate from linearity only in that some one rule may be permitted to apply at more than one point in the ordering, and secondly, any rule which applies at more than one point

must be such that no other rule must apply after its last application. This second requirement follows from the fact that persistent rules are always present at depth 1. Chafe's proposal further makes the claim, with which I can associate no empirical consequences, that violations of linearity are always a result of a particular rule's having a peculiar status (that of persistent rule) as opposed to all of the other rules of the grammar, rather than being a consequence of the general properties of the ordering relation, as is assumed by a theory that attempts to make crucial use of the notion of unmarked order in defining such violations.

In fact, both of the particular claims Chafe makes about 'paradoxical' orderings seem false, on the basis of examples that have been considered here. The claim that the solution must be obtainable by allowing a rule to appear at more than one point is falsified by the Sanskrit example considered above in section V.4: since the two rules GL and DaC must apply in different relative orders depending on the applicability of the rule BL, one or the other must be 'persistent' in Chafe's framework. In that case, the solution must be obtainable either by the ordering GL-BL-DaC-GL or by the ordering DaC-GL-BL-DaC. The reader can satisfy himself, however, that neither of these orders will give the correct solutions.

The requirement that persistent rules must appear at depth 1 is falsified by the example of u-umlaut and truncation that was considered in chapter II above. It will be recalled that u-umlaut had to apply both before and after truncation; it must, then, be a persistent rule. But it cannot appear at depth 1; the rule of u-epenthesis, which applies before e.g. the nominative singular ending /+r/, produces new instances of u to which the u-umlaut rule must not apply. If u-umlaut appears at depth 1, however, u-epenthesis must appear at at least depth 2, since their relative ordering

makes a difference, and hence they cannot appear on the same level. In this case, there will be no way to prevent the u's produced by epenthesis from causing umlaut. Indeed, one of Chafe's own examples falsifies his claim: in his French example, the rule of schwa-insertion (6.3a) must be persistent, and hence must appear at depth 1; but it is a fact that the schwas inserted by this rule may be truncated by the elision rule (not discussed here, but familiar) which applies before words beginning with a vowel. Although there may be no particular reason to prohibit the schwa-insertion rule from appearing at the depth of the elision rule, as well as before it (in at least the two places required by Chafe), this step has no empirical consequences whatever, and seriously reduces the content of the claim that persistent rules appear at depth 1 at all. We must conclude, then, that while Chafe has provided evidence in support of a relaxation of the requirement of linearizability, the particular features of his proposal are not supported by the evidence, and that the theory espoused here provides a closer approximation to the truth.

## Notes

<sup>1</sup> Among other things, the non-occurrence of the sequences ye, we, yo, wo can be predicted in terms of the rules (independently necessary) which would reduce the sequences underlying such surface forms to e or o.

<sup>2</sup> Actually, each of these vowels has two different shapes, divided into harmony classes. This harmony, described by Callow, is disregarded here.

<sup>3</sup> Observe, however, that the alternation pia/pe could be derived not only from the stem /pia/, but also from the hypothetical stem /pai/; given /pai+a/ and /pai+i/, a metathesis rule restricted from applying to sequences of identical vowels (regardless of their quality) will apply to the singular form (only), giving /pia+a/, with the plural remaining /pai+i/, which will become [pia] and [pe]. There is no particular reason to choose either of the possible stems, /pai/ and /pia/ rather than the other (at least on the basis of this data), but if the restriction on metathesis is generalized to exclude all identical sequences (thus becoming identical with the SD for truncation), as argued here, we see that the two possibilities would lead to homophonous forms. This would account for the absence of nouns like a hypothetical \*pia/pye, which would result from the stem /pai/ under the exact set of rules in SPE.

<sup>4</sup> In this form, h is from underlying /gh/.

<sup>5</sup> For the use of this notion in phonology, cf. Lightner (1968).

<sup>6</sup> This is the section containing the rule referred to here as AM.

<sup>7</sup> Whitney, 1889: sec. 160b.

<sup>8</sup> We use the symbols [ $k_1$ ,  $g_1$ , and  $l_1$ ] to represent the fronted segments corresponding to [k, g, l]. Some authors have used [ $\lambda$ ] for the back  $\underline{l}$ ,  $\underline{g}$  for the back  $\underline{k}$ , etc. These transcriptions have been consistently altered to conform to the usage here.

<sup>9</sup> This simultaneously progressive and regressive character of the velar harmony rule has led some authors to the radical claim that this kind of process is not a rule of phonology at all, but rather a process whereby an abstract (arbitrary) 'root marker' feature is distributed over the whole word, and then used to condition a rule which in effect realizes the ad hoc feature as a phonological one. The development of a theory of neighborhood rules, as suggested in section IV.4 above, obviates this move.

<sup>10</sup> The feature [- strident] in this rule is intended to exclude the segments [š, ž, č, ʃ] from the affected class; it may be, however, that these segments are actually velarized in this environment, in which case they can be included. My knowledge of Turkish phonetics is inadequate to decide this issue.

<sup>11</sup> The various harmony rules discussed here seem to apply across morpheme boundaries only. Other rules with the same effects apply within morphemes as rules of lexical structure (or 'morpheme-structure rules'); the relation between the two has been glossed over at a number of points in this

work. A detailed investigation of some aspects of this relation is being undertaken by M. Brame in unpublished work.

<sup>12</sup> There is also at least one word that has a final front vowel, but is said to take back vowel suffixes: Utarid 'Mercury (planet)'. I have no idea of how this form should be handled.

<sup>13</sup> A morpheme structure rule to this effect will still be needed, however, as well as one specifying all suffix vowels as [- round].

<sup>14</sup> The environment '/ + \_\_\_\_ +' is my interpretation of Chafe's use of the cover symbol 'V<sub>t</sub>' for thematic vowels.

## Chapter VI - Local Ordering and Phonological Theory

VI.1 Some implications. The theory of local ordering which has been defended in this work is based on two principles: first, that ordering relations between rules are pairwise, and that they are independent of one another; and secondly, that the notion of marked vs. unmarked order is central to the definition of orderings. These principles, and the theory of ordering relationships which is based on them, seem to relate to other areas of linguistic theory.

For one thing, such a theory appears to offer a more plausible basis for a model of linguistic acquisition and performance than does a theory based on linear order. In the process of learning a language, a child presumably develops a grammar with certain rules; when he incorporates forms to which more than one of these rules apply, he learns ordering relationships. A theory of grammar based on marked and unmarked orders assumes that the rules are used wherever possible, and extended to as many new classes of forms as possible. On such a theory, it is only in cases where the applicability of one of the rules must be restricted that an ordering relationship must be learned. Without the notion of unmarked order, any ordering relationship ought to be just as learnable as any other, and there is no reason to expect ordering relationships to conform to a principle of maximizing applicability. Yet we know (cf. Kiparsky, 1968b, for discussion) that the direction of historical change is toward a reduction in the number of marked orderings; the basis of the explanatory power of 'analogy' resides in the extension of rule applicability to new classes of forms; and even a cursory examination of any grammar will show that more ordering relations are unmarked than marked. This suggests that the notion of unmarked order

is part of the principle of application of grammars, and not simply a device for evaluating already given grammars.

The language-learning child has to learn certain ordering relationships between rules of his grammar; the theory of linear ordering asserts that he constrains these relations, requiring them to conform to at least one linear ordering, and furthermore that he defines ordering relations even between pairs of rules that do not interact. If grammars in fact conformed to these conditions, this would be evidence that the acquisition procedure is in fact so constrained; the evidence cited above, however, seems to indicate that this is not the case, and that the acquisition device does not operate under these conditions. Since such a requirement would be an arbitrary one, unrelated to any other known property of the mind, it seems a gain in naturalness to be able to do without it.

Further, if the individual ordering relationships in the grammar are independent facts, one would expect it to be possible for forms in the language to be exceptional with respect to one or more of these restrictions, just as they can be exceptional with respect to the rules themselves, without changing the rest of the grammar. In fact, as Bailey (1968) has recently argued, it is necessary to include in the theory of exceptionality statements such as "where rules  $R_1$  and  $R_2$  normally apply in a marked order, in the form  $F$  they apply in unmarked order." An example cited by Bailey is the difference between, say, relation and equation in English. These are derived, in the SPE analysis of English, from /rE=1At+ion/ and /E=kWAt+ion/, respectively; in both, processes of spirantization of the stem-final t apply. There is another rule in English which voices certain intervocalic spirants; this rule is in a marked order with respect to the spirantization process, however, and spirants created by the latter are not generally voiced by the former. This



accounts for the voicelessness of the spirant in relation. Equation, however, has a voiced spirant; this could be accounted for by assuming that it is marked as exceptional in that spirantization and spirant voicing apply in their unmarked order in the derivation of this form. Further discussion of the role of this kind of exceptionality in English can be found in Carlson (1969); in chapter II above we saw some cases in which a similar sort of exceptionality was required in Old Norse. The existence of such exceptions indicates that ordering relations have the same sort of status as rules in the grammar, rather than being overall conditions of application; one cannot readily envision the possibility of a dictionary item which was exceptional in that the principle of disjunctive order did not apply to rules affecting this form.

VI.2 The status of 'marking conventions.' In SPE, Chomsky and Halle confronted the problem of providing a substantive interpretation for the feature system and other devices of phonological theory. In formulating their proposal for the incorporation into linguistic theory of a theory of marking conventions, they hoped to deal with the problem of specifying the extent to which particular feature complexes are 'natural' in a more satisfactory way than the device of counting features provides. For example, it is more natural (on the basis of one's impressions from the languages of the world) for vowels to agree in rounding and backness (if they are [- low]) than for them to disagree in these features; yet a lexical entry containing the features [-low, +back, -round, +high] is no more complex in terms of number of features than one containing [-low, +back, +round, +high]; i.e., there is no difference between [u] and [ɯ] in these terms. Chomsky and Halle proposed to introduce a set of interpretive principles, called marking

conventions, which would evaluate the various values for the features in a lexical entry as relatively marked (unnatural) or unmarked (natural). In the example just given, for instance, they would say that if the feature [ $\pm$  round] for a given segment agrees with the feature [ $\pm$  back] (in the presence of the feature [- low]), this value is unmarked. The feature complexes whose evaluation is relevant for determining the relative complexities of alternative formulations of part of a grammar, then, are not simply matrices of +'s and -'s, but rather these matrices as interpreted by the marking conventions: if a feature value is unmarked, it adds nothing to the complexity of the matrix, while if it is marked, it adds one unit of complexity. The marking conventions, then, served as an evaluation function for the feature complexes appearing in lexical entries.

In addition to this, however, it was hoped that the marking conventions could contribute to the definition of the notion of 'natural rule'. For example, just as the segment [s] is more natural than [θ], so if a rule turns [t] into a spirant, it is more natural for it to turn it into [s] than into [θ]. This difference is not reflected in the relative complexity of the rules as stated in features, however; both rules must make the t [+ continuant], while the rule turning it into s must additionally specify it as [+ strident] and [+ distributed]. Chomsky and Halle proposed to remedy this problem by permitting the marking conventions to have an interpretation as phonological rules; whenever a feature or features are changed, other features are changed by the marking conditions to take their unmarked values, unless the rule explicitly states otherwise. The rule spirantizing t to s, then, need only make it [+ continuant]; the fact that the unmarked values for the features of stridency and distributedness in dental spirants are [+strid, +dist] will result in the automatic insertion of these features.

The rule producing θ, however, will have to be stated so as to require the result [-strid, -dist]; thus, this latter rule will cost two more features than the rule producing s.

The characteristics of marking conventions then, are, first universality; that is, since they are to be part of linguistic metatheory, rather than part of the grammars of particular languages, they must be the same for all languages. Secondly, they are 'rules' that perform the same operations on lexical entries (that is, as 'morpheme structure rules') and on derived representations (that is, as 'phonological rules'). Thirdly, since the marking conventions can apply at any point where the conditions for their operation are met, they are not constrained to a single appearance in the ordering of rules.

While it is clear that the problems to which the theory of marking conventions is designed to provide an answer are very real, it is not clear that this particular device is the correct approach. Consider the issue of universality, for example: surely one of the most widely applicable 'unmarked states' is for a nasal to agree in point of articulation with that of an adjacent obstruent. Yet different languages have different forms of this process; in Pame (cf. section I.2 above), the nasal assimilation process is conditioned by an obstruent on either side of the affected nasal (i.e., it is a neighborhood rule), while in many languages (e.g., English) only a following obstruent conditions assimilation. Since both are surely 'nasal assimilation' processes, it seems necessary to accept language-particular formulations of very general processes.

For another example, consider voicing assimilation processes. No set of marking conventions that fails to capture the fact that homogeneity in the feature [ $\pm$  voice] over a consonant cluster is the natural state can

possibly be called adequate; yet various languages have distinct realizations of this fact. Consider the operation of this process in the proto-Germanic language just after the addition to the grammar of the rules expressing Grimm's law.<sup>1</sup> At this stage, the verb underlying, e.g., Gothic skapjan had an underlying stem /skab-/; before a voiceless ending such as that underlying Go. ga-skafts, the b was devoiced to p, and then this p was affected to give the p of skapjan. Thus, a rule of voicing assimilation existed which applied before the Grimm's law rule(s). In a form like Go. asts 'bough', from underlying \*ozdos, however, the voicing assimilation process must apply after Grimm's law has changed d to t. Hence, the voicing assimilation process would have to have the ordering properties of a marking convention (on the SPE theory). In English, however, there is a voicing assimilation process which operates not regressively, like the one just discussed, but rather progressively. In English inflectional endings (as well as the phonologically reduced forms of the auxiliary verb is to 's'),<sup>2</sup> the voicing of an obstruent is assimilated to that of a preceding one:<sup>3</sup>

$$2.1 \quad [+ \text{obst}] \rightarrow [- \text{voice}] / \begin{bmatrix} + \text{obst} \\ - \text{voice} \end{bmatrix} \# \_\_\_\_\_\_ \#$$

Since both this progressive rule and the regressive assimilation discussed above should surely be considered representatives of the universal tendency to voicing assimilation, it seems we are once again led to language particular formulations of 'marking conventions', or at least of the kind of facts they were designed to handle.

In connection with the claim that marking conventions have reflexes both as morpheme structure rules and as phonological rules, recall the vowel harmony processes in Turkish. These too had reflexes in both parts of the grammar; the fact that the morpheme structure rule which corresponds to a

given harmonic constraint is a separate rule from the phonological process applying across morpheme boundaries to achieve the same ends is shown by the fact that they have distinct classes of exceptions; a form like ziyaret 'visit' is exceptional with respect to its morpheme-internal vowels in the feature of [ $\pm$  back], but regular in the forms of the affixes it takes; while kabahat 'fault' is regular morpheme-internally, but irregular with respect to the affixes it takes: cf. kabahatti 'his fault'. Since the Turkish backness harmony rule is surely a language particular rule and not a marking convention, we see that such rules have the same properties of applying both as MS rules and as phonological rules. Similar arguments have been given by Kiparsky (1968a) for other languages; as was mentioned above, the problem has recently been considered by Brame, as well.

With respect to the special ordering properties posited for marking conventions applying as linking rules, the examples already discussed in this work in support of the theory of local ordering have shown that many language particular phonological rules can also appear at more than one point in the sequence of rules. Indeed, the special ordering necessary for marking conventions is simply the absence of any marked orderings; in this case, they will apply in unmarked order with respect to all other rules, and hence any rule that produces representations to which they could apply will 'feed' such a rule. This absence of any marked orderings corresponds well to the presumed generality of such rules.

We have seen, therefore, that while a set of interpretive conventions to express the relative naturalness of feature matrices is certainly necessary, the extension of these conventions to a use as linking rules to account for the naturalness of phonological rules is suspect. In fact, all of the properties which were attributed to such rules are seen to be characteristic

of at least some phonological rules that are so obviously language particular that no one would want to call them marking conventions. Further, the marking conventions themselves seem to require language particular formulation. Much of the role of marking conventions, then, seems to be just that of very general phonological rules.

The function that marking conventions were supposed to perform by functioning as linking rules was to provide a theory of rule naturalness; at least part of this function can be performed by a natural extension of the notion of marking conventions as an evaluation function. As suggested (in essence) by Postal (1968), the marking conventions can be used to evaluate the complexity, not only of feature complexes which appear in (fully specified) lexical entries, but also of the feature complexes which occur in (partially specified) phonological rules. That is, the two rules mentioned above could be formulated approximately as:

$$\begin{array}{ll}
 2.2 \text{ a. } (t \rightarrow s) & \begin{bmatrix} + \text{ cor} \\ - \text{ cont} \\ - \text{ nas} \end{bmatrix} \rightarrow \begin{bmatrix} + \text{ cont} \\ u \text{ strid} \\ u \text{ dist} \end{bmatrix} \\
 \text{b. } (t \rightarrow \theta) & \begin{bmatrix} + \text{ cor} \\ - \text{ cont} \\ - \text{ nas} \end{bmatrix} \rightarrow \begin{bmatrix} + \text{ cont} \\ m \text{ strid} \\ m \text{ dist} \end{bmatrix}
 \end{array}$$

While other features of these rules would presumably be evaluated as marked or unmarked, these specifications would presumably be the same for both rules; the naturalness of 2.2a as opposed to 2.2b is reflected in the fact that the former has two unmarked (and hence cost-free) specifications where the latter has marked (and hence costly) specifications.

Only a portion of the problem of rule evaluation is amenable to this sort of treatment, however. The voicing assimilation in segments adjacent to consonants affected by Grimm's law (as in \*ozdos  $\rightarrow$  asts) is clearly a

separate fact from the Grimm's law change itself; while it might be possible to incorporate all of the possible side effects of a rule into the rule itself, this would leave the unity of such processes unexpressed. In fact, some sort of theory of naturalness of rules is necessary which allows rules to be evaluated as less costly by virtue of their being the reflexes in a particular language of very general processes of assimilation, etc. Such a theory is essayed, in part, by Charles Kisseberth (forthcoming); no effort will be made here to indicate any properties of such a theory other than the fact that it should be a principle by which rules as written are evaluated, rather than a restructuring of the way the rules are to be written, as is the case with the theory of 'linking' marking conventions. It is simply not the case that formal complexity can be made to reflect naturalness directly, since the structure of the statement of a rule is dictated by other factors. Phonological features have (by the naturalness criteria proposed in Postal, 1968) a direct interpretation as phonetic features (with articulatory and acoustic reflexes), and, given a phonological generalization, its statement in features should be inflexibly determined by the phonetic framework. It is of course possible to deform the notation in such a way as to reflect any sort of property desired, but not without impairing the naturalness of the relation between its statement and its phonetic interpretation. The feature framework (and thereby the formulation of rules) is a fact independent of the naturalness of particular statements in it -- a given on which an evaluation measure operates. Markedness is a matter of the evaluation of grammars, not of their statement; hence it belongs in the evaluation measure as an interpretive principle. Just as the device of braces or parentheses has consequences for the evaluation of rules which are relatable in certain ways without altering the statement of the subrules it

relates, so also the theory of natural phonology should evaluate rules without altering them.

In addition to the naturalness of the feature complexes in terms of which a rule is stated, and the extent to which a rule 'participates in' or 'realizes' a universal phonetic process, it is necessary to capture the generalization which is inherent in the relation between a morpheme structure condition and a phonological rule with the same effect. Actually, as has recently been argued in an interesting paper by Kisseberth (1969), this is an aspect of the fact that several different rules in a language may 'conspire' to eliminate some tactically impossible configuration, or to produce only configurations of a certain type, etc. In Kisseberth's example, a number of rules in Yawelmani operate to insure the elimination of all sequences of three or more consonants (or two at # boundary). A morpheme structure rule, rules that delete consonants under certain conditions, and rules that insert epenthetic vowels to break up other sequences of three consonants all have this effect. There is surely a single generalization here: this grammar should be evaluated as more natural than one with a set of individual rules, each of the same complexity as one of the Yawelmani rules, but which are unrelated to one another. Phonological theory has only begun to realize the existence of this sort of problem, let alone produce an adequate theory to deal with it. In particular, it is necessary to come to grips with the problem of delimiting those of the vast number of 'conspiracies' that could be found in a grammar that are significant from those that are purely accidental. Until this is done, the notion of conspiracy remains a vacuous, if suggestive one.

The contribution of the theory of local ordering to this issue, and to the others raised above, is simply to provide a framework within which



the phonological generalizations of a language can be stated adequately, a clear precondition to any evaluation. By presenting evidence that the correct solutions to phonological problems do not in general conform to a condition of linearity, this work has removed the necessity of deforming the statement of particular rules so as to permit them to be arranged linearly. We have seen several instances above (e.g., Zwicky's and Kiparsky's analyses of Grassmann's law in Sanskrit) where the need to maintain this condition led to incorrect and unnatural statements. By freeing phonology from such a constraint, it becomes possible to obtain a clearer idea of what processes are natural in natural language.

## Notes

<sup>1</sup> For details, cf. Streitberg, 1963, and Wright, 1954.

<sup>2</sup> For discussion, cf. Luelsdorf, 1969.

<sup>3</sup> Morris Halle (personal communication) has pointed out to me that alternations like five/fifth and leave/left are exceptions to the process described here. I suspect that a case could be made for representing them with underlying voiceless segments, followed by a vowel which is deleted in the derived forms. The intervocalic f, then, will be voiced by a natural extension of the English rule of s-voicing. Thus, five = /fIfe/, leave = /lEfe/; fifth = /fIfe + θ/, left = /lEfe+Vd/ → /lEf+Vd/ → /lEf+d/ → /lEf+t/. I wouldn't want to push this, however.

## Bibliography

- Anderson, Stephen R. (1968). "The Faroese Vowel System and the Faroese Verschärfung." in Quarterly Progress Report of the Research Laboratory of Electronics, 90:228-40.
- \_\_\_\_\_ (1969). "An Outline of the Phonology of Modern Icelandic Vowels." Foundations of Language V:53-72.
- Bach, Emmon (1968). "Two Proposals Concerning the Simplicity Metric in Phonology." Glossa, 2:128-49.
- Bailey, Charles-James N. (1968). "An Untested Idea on Lexical Exceptions to the Regular Ordering of the Phonological Rules of a Language." ERIC/Pegs paper number 25.
- Benediktsson, Hreinn (1958). The Vowel System of Icelandic: Its Structure and Development. Unpublished Ph.D. Dissertation, Harvard University.
- \_\_\_\_\_ (1959). "The Vowel System of Icelandic: A Survey of its History." Word, 15:409-31.
- \_\_\_\_\_ (1960). "Phonemic Neutralization and Inaccurate Rhymes." Acta Philologica Scandinavica, 26:1-18.
- \_\_\_\_\_ (1963a). "Some Aspects of Nordic Umlaut and Breaking." Language, 39:409-31.
- \_\_\_\_\_ (1963b). "The Non-Uniqueness of Phonemic Solutions: Quantity and Stress in Icelandic." Phonetica, 10:133-53.
- \_\_\_\_\_ (1968). "Indirect Changes of Phonological Structure: Nordic Vowel Quality." Acta Linguistica Hafniensa, XI:31-65.
- \_\_\_\_\_ (1969). "Aspects of Historical Phonology." Paper read at the International Conference on Nordic and General Linguistics, Reykjavík, Iceland; July, 1969.
- Bever, Thomas G. (1967). Leonard Bloomfield and the Phonology of the Menomini Language. Unpublished Ph.D. Dissertation, Massachusetts Institute of Technology.
- Bjerrum, Marie (1949). "An Outline of the Faroese Vowel System." Travaux du Cercle Linguistique de Copenhague, 5:235-43.
- Bloomfield, Leonard (1939). "Menomini Morphophonemics." Travaux du Cercle Linguistique de Prague, 8:105-15.
- \_\_\_\_\_ (1962). The Menomini Language. New Haven: Yale University Press.

- Bright, William (1966). "Dravidian Metaphony." Language, 42:311-22.
- Brunner, Karl (1963). An Outline of Middle English Grammar. Cambridge: Harvard University Press.
- Burrow, T. (1959). The Sanskrit Language. London: Faber & Faber.
- Callow, J. C. (1965). "Kasem Nominals -- A Study in Analyses." Journal of West African Languages, 2:29-36.
- \_\_\_\_\_. (1968). "A Heirarchical Study of Neutralization in Kasem." Journal of Linguistics, 4:33-46.
- Carlson, Barry F. (1969). "Unmarked Order and Lexical Exceptions." University of Hawaii Working Papers in Linguistics, 6:205-12.
- Chafe, Wallace L. (1968). "The Ordering of Phonological Rules." International Journal of American Linguistics, 34:115-36.
- Chapman, Kenneth (1962). Icelandic-Norwegian Linguistic Relationships. Norsk Tidsskrift for Sprogvidenskap, Suppl. Bind VII.
- Chomsky, Noam (1957). Syntactic Structures. The Hague: Mouton & Co.
- \_\_\_\_\_. (1964). Current Issues in Linguistic Theory. The Hague: Mouton & Co.
- \_\_\_\_\_. (1967). "Some General Properties of Phonological Rules." Language, 43:102-28.
- \_\_\_\_\_, and Morris Halle (1967). "Some Controversial Questions in Phonological Theory." Journal of Linguistics, 1:97-138.
- \_\_\_\_\_. (1968). The Sound Pattern of English. New York: Harper and Row.
- \_\_\_\_\_, and Fred Lukoff (1956). "On Accent and Juncture in English." in Halle, Lunt and MacLean (eds), For Roman Jakobson, 65-80. The Hague: Mouton & Co.
- Dave, Trimbaklal N. (1935). A Study of the Gujarati Language in the 16th Century. London: Royal Asiatic Society.
- Einarsson, Stéfan (1945). Icelandic. Baltimore: Johns Hopkins Press.
- Elcock, W.D. (1960). The Romance Languages. London: Faber & Faber.
- Fleuriot, Leon (1964). Le Vieux Breton. Paris: Librairie C. Klincksieck.
- Foster, Joseph F. (1968). "On Vowel Harmony and Concomitant Phenomena in Osman Turkish." Unpublished ditto, University of Illinois.

- Gibson, Lorna (1956). "Pame (Otomi) Phonemics and Morphophonemics." International Journal of American Linguistics, 22:242-65.
- Glendening, P.J.T. (1961). Teach Yourself Icelandic. London: The English Universities Press.
- Gonda, Jan (1966). A Concise Elementary Grammar of the Sanskrit Language. University, Ala.: The University of Alabama Press.
- Gordon, E.V. (1957). An Introduction to Old Norse. London: Oxford University Press.
- Gutenbrunner, Siegfried (1951). Historische Laut- und Formenlehre des Altisländischen. Heidelberg: Carl Winter Verlag.
- Høegstad, Marius (1917). Vestnorske Maalføre fyre 1350. Vol. II: Sudvestlandsk. Kristiania: I Kommission hos Jacob Dybwad.
- Halle, Morris (1958). "On the Bases of Phonology." in Fodor and Katz (eds), The Structure of Language. Englewood Cliffs, N.J.: Prentice Hall.
- \_\_\_\_\_ (1959). The Sound Pattern of Russian. The Hague: Mouton & Co.
- \_\_\_\_\_ (1962). "Phonology in Generative Grammar." in Fodor and Katz (eds), The Structure of Language. Englewood Cliffs, N.J.: Prentice Hall.
- \_\_\_\_\_ (1967). "On Meter and Prosody." Unpublished ditto, Massachusetts Institute of Technology.
- \_\_\_\_\_, and S. Jay Keyser (1966). "Chaucer and the Study of Prosody." College English, . 82-219.
- \_\_\_\_\_ (forthcoming). The Evolution of English Stress. New York: Harper and Row.
- Halpern, A.M. (1946). "Yuma I: Phonemics." International Journal of American Linguistics, 12:25-33.
- Hammershaimb, Venceslaus U. (1891). Færøsk Anthologi. Copenhagen: S.L. Møllers Bogtrykkeri.
- Hardie, D.W.F. (1948). A Handbook of Modern Breton. Cardiff: University of Wales Press.
- Haugen, Einar (1949). "The Unstressed Vowels of Old Icelandic." Norsk Tidsskrift for Sprogvidenskap, 15:384ff.
- \_\_\_\_\_ (1950). The First Grammatical Treatise: The Earliest Germanic Phonology. Language Monograph number 25.
- \_\_\_\_\_ (1958). "The Phonemics of Modern Icelandic." Language, 34:55-88.

- Heusler, Andreas (1964). Altisländisches Elementarbuch. Heidelberg: Carl Winter Universitätsverlag.
- Jakobson, Roman (1952). "Studies in Comparative Slavic Metrics." in Selected Writings, vol. IV. The Hague: Mouton & Co.
- \_\_\_\_\_ (1963). "On the So-called Vowel Alliteration in Germanic Verse." Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung, 16:85-92.
- Jónsson, Finnur (1929). Carmina Scaldica: Udvalg af Norske og Islandske Skaldekvad. Copenhagen: G.E.C. Gads Forlag.
- Jónsson, Snæbjörn (1927). A Primer of Modern Icelandic. London: Oxford University Press.
- Kahle, Bernhard (1892). Die Sprache der Skalden. Strassburg: Karl J. Trübner Verlag.
- Keller, Kathryn (1959). "The Phonemes of Chontal (Mayan)." International Journal of American Linguistics, 25:44-53.
- Kelley, Gerald B. (1963). "Telugu Vowel Phonemes and Vowel Sandhi." Journal of the American Oriental Society, 83:67-73.
- Keyser, S.J. (1968). "Old English Prosody." in Anderson, Jackendoff and Keyser, Studies in Transformational Grammar and Related Topics. AFCRL final report, Brandeis University.
- Kiparsky, Paul (1965). Phonological Change. Unpublished Ph.D. Dissertation, Massachusetts Institute of Technology.
- \_\_\_\_\_ (1968a). "How Abstract is Phonology?" Unpublished ditto, Massachusetts Institute of Technology.
- \_\_\_\_\_ (1968b). "Linguistic Universals and Linguistic Change." in Bach and Harms (eds), Universals in Linguistic Theory. New York: Holt, Rinehart and Winston.
- \_\_\_\_\_ (1968c). "Metrics and Morphophonemics in the Kalevala." in Studies Presented to Roman Jakobson by his Students. Cambridge, Mass.: Slavica Press.
- Kisseberth, Charles W. (1969). "On a Conspiracy in Yawelmani." Unpublished ms., University of Illinois.
- \_\_\_\_\_ (forthcoming). Theoretical Implications of Yawelmani Phonology. University of Illinois Ph.D. Dissertation.
- Kock, Axel (1902). "Der I-Umlaut von e in den Altnordischen Sprachen." Beiträge zur Geschichte der Deutschen Sprache. 27:166-190.

- Kreider, Herman H. (1954). Essentials of Modern Turkish. Washington: The Middle East Institute.
- Krenn, E. (1940). Föroyische Sprachlehre. Heidelberg: Carl Winters Universitätsbuchhandlung.
- Kučera, Henry (1958). "Inquiry into Co-existent Phonemic Systems in Slavic Languages." in American Contributions to the 4th International Conference of Slavists. The Hague: Mouton & Co.
- Ladefoged, Peter (1968). A Phonetic Study of West African Languages<sup>2</sup>. West African Languages Monographs. Cambridge: Cambridge University Press.
- Lamb, Sydney M. (1966). "Prolegomena to a Theory of Phonology." Language, 42:536-73.
- Lane, George S. (1946). (Review of Einarsson (1945)). Language, 22:253-5.
- Lees, Robert B. (1961). The Phonology of Modern Standard Turkish. Indiana University publication number 6 in the Ural-Altaic Series. The Hague: Mouton & Co.
- \_\_\_\_\_ (1967). "Turkish Harmony and the Phonological Description of Assimilation." Türk Dili Araştırmaları Yıllığı-Belleten 1966'dan ayırışım. Ankara: Ankara Üniversitesi Basımevi.
- Lightner, Theodore M. (1968). "On the Use of Minor Rules in Russian Phonology." Journal of Linguistics, 4:69-72.
- Lockwood, W.B. (1955). An Introduction to Modern Faroese. Copenhagen: Ejnar Munksgaard.
- Luelsdorff, Phillip (1969). "On the Phonology of English Inflection." Glossa, 3:
- McCawley, James D. (1964). "Morphophonemics of the Finnish Noun." Unpublished ditto, Massachusetts Institute of Technology.
- \_\_\_\_\_ (1966). "Revised Version of Finnish Rules." Unpublished ditto, University of Chicago.
- \_\_\_\_\_ (1967). "Further Revisions of Finnish Rules." Unpublished ditto, University of Chicago.
- \_\_\_\_\_ (1968). "A Note on Faroese Vowels." Glossa, 2:11-16.
- Moore, Samuel (1951). Historical Outline of English Sounds and Inflections. Ann Arbor, Michigan: George Wahr Publ.
- Noreen, Adolf (1923). Altisländische und Altnorwegische Grammatik. Halle (Saale): Max Niemeyer.

- O'Neil, Wayne A. (1964). "Faroese Vowel Morphophonemics." Language, 40:366-71.
- Poppe, Nicholas (1965). Introduction to Altaic Linguistics. Wiesbaden: Otto Harrassowitz.
- Postal, Paul (1968). Aspects of Phonological Theory. New York: Harper and Row.
- Rask, Rasmus (1830). Die Verslehre der Isländer. Berlin: G. Reimer.
- Reichardt, Konstantin (1928). Skalden des 9 u. 10 Jahrhunderts. Leipzig: Mayer und Müller.
- Rischel, Jørgen (1968). "Diphthongization in Faroese." Acta Linguistica Hafniensa, XI:89-118.
- Roe, H.A. (1965). Verschärfung in Faroese. Unpublished Ph.D. Dissertation, Harvard University.
- Sievers, Eduard (1878,9). Beiträge zur Skaldenmetrik. Halle: E. Karras.
- Sigurd, Bengt (1966). "Generative Grammar and Historical Linguistics." Acta Linguistica Hafniensa, X:35-48.
- Staal, J.F. (1968). "The Ordering of Rules in Panini's Grammar." Paper read at Winter, 1968 meeting of the Linguistic Society of America, New York City.
- Streitberg, W. (1963). Urgermanische Grammatik<sup>3</sup>. Heidelberg: Carl Winter Universitätsverlag.
- Sturluson, Snorri (ed. 1954). Edda Snorra Sturlusonar. Akureyri: Prentverk Odds Björnssonar.
- Sturtevant, Albert M. (1927). "Zum Lautwandel a : ió : ia im Altnordischen." Germanic Review, II:334-52.
- \_\_\_\_\_. (1953). "The Old Norse Labialization of a, Due to an ö of the End Syllable in the Adjectival Type \*skall-öttr > sköll-öttr 'bald'." Modern Language Notes, Dec. 1953:542-3.
- Swift, Lloyd B. (1963). A Reference Grammar of Modern Turkish. Indiana University Publication number 19 of the Ural-Altaic Series. The Hague: Mouton & Co.
- Thurneysen, Rudolf (1961). A Grammar of Old Irish. Dublin: Dublin Institute of Advanced Studies.
- Valfells, Sigrid (1967). Umlaut-Alternations in Modern Icelandic. Unpublished Ph.D. Dissertation, Harvard University.



- Voegelin, C.F. (1936). "Tübatulabal Grammar." in University of California Publications in American Archaeology and Ethnography, vol. 34.
- Werner, Otmar (1964). "Die Erforschung der Färingischen Sprache. Ein Bericht über Stand und Aufgaben." Orbis, XIII: 481-544.
- \_\_\_\_\_ (1968). "Welche Stufen Phonematischer Reduktion sind für die Dialektgeographie Sinnvoll? Das Vokalsystem des Färöischen." Zeitschrift für Mundartforschung, 4:861-870.
- Whitney, William D. (1889). Sanskrit Grammar. Cambridge: Harvard University Press (republished, 1964).
- Wolfe, Patricia (1968). "Exchange Rules and the Impairment of Intelligibility". Paper read at Winter, 1968 meeting of the Linguistic Society of America, New York City.
- Wright, Joseph (1928). An Elementary Middle English Grammar. London: Oxford University Press.
- \_\_\_\_\_ (1954). Grammar of the Gothic Language.<sup>2</sup> London: Oxford University Press.
- Zeps, Valdis J. (1963). "The Meter of the So-called Trochaic Latvian Folk-songs." International Journal of Slavic Linguistics and Poetics, VII: 123-8.
- Zwicky, Arnold M. (1965). Topics in Sanskrit Phonology. Unpublished Ph.D. Dissertation, Massachusetts Institute of Technology.
- \_\_\_\_\_ (1968). "Umlaut and Noun Plurals in German." Studia Grammatica, VIII:

## Biographical Note

The author was born on 3 August, 1943, in Madison, Wisconsin, and attended school there and in Orlando, Florida. In September of 1960 he entered the University of Chicago, where he studied mathematics. He received the twelfth grade certificate in June of 1961. Upon leaving school in January, 1963, he spent some time as a computer programmer, an occupation he found even more distasteful than mathematics. Accordingly, he enrolled in classes at the Illinois Institute of Technology, later becoming a full-time student in linguistics. In June, 1966, he was awarded IIT's first B.S. degree in linguistics. After attending the 1966 LSA Summer Institute at UCLA, he entered MIT. During the next three years, he held first an NIH traineeship and later an NDEA Title IV fellowship. At various points in his graduate career, he participated in research at Brandeis University under S. Jay Keyser, at the Harvard University Computation Laboratory under Susumu Kuno, and at the Language Research Foundation under Bruce Fraser. In the summer of 1968 he visited the Faroe Islands and Iceland with the help of a grant from the Thor Thors fund of the American-Scandinavian Foundation. Upon receipt of his degree from MIT, he will take a position as Assistant Professor of Linguistics at Harvard University.